MontCAS
(Montana Comprehensive Assessment System)
English Language
Proficiency Assessment

MontCAS English Language Proficiency (ELP) Assessment

Technical Report

2008-2009

Copyright © 2010, held by the Montana Office of Public Instruction. All rights reserved.

Printed in the U.S.A

Table of Contents

	Page
1. Purpose of the Technical Report	1
2. Description of the MontCAS ELP	1
2.1 Purpose of the MontCAS ELP	1
2.2 Past and Present MontCAS ELP Forms	2
2.3 Structure of the MontCAS ELP	3
2.4 Alignment of the MontCAS ELP	7
3. MontCAS ELP 2008-2009 Administration	7
3.1 Testing Window	7
3.2 Assessment Training	7
3.3 Test Administrator Scripts	7
3.4 Listening Test Administration	8
3.5 Setting for the Test	8
3.6 Timing	8
3.7 Prompting and Repeating Test Information	8
3.8 Testing Absentees	9
3.9 Testing Accommodations	9
4. MontCAS ELP 2008-2009 Test Security	10
4.1 Bar-Coding and Return of Secure Materials	11
4.2 Storage and Shredding of Secure Materials	11
5. MontCAS ELP 2008-2009 Scoring and Reporting	11
5.1 Scoring of Multiple-Choice Items	11
5.2 Writing Checklist	11
5.3 Scoring of Constructed-Response Items	11
5.4 Reporting	14

Table of Contents (Continued)

	Page
6. MontCAS ELP 2008-2009 Item Analyses	16
7. Scaling and Equating of the MontCAS ELP	18
8. Reliability of the MontCAS ELP	23
9. Validity of the MontCAS ELP	28
9.1 Content-related Validity	28
9.2 Construct and Criterion-related Validity	28
10. MontCAS ELP Performance by Year	31
11. References	40
12. Appendices	
A. MontCAS ELP Item Difficulty and Discrimination Data	41
Table A.1. Grade K: Form A	41
Table A.2. Grades 1-2: Form B1	43
Table A.3. Grades 1-2: Form B2	45
Table A.4. Grades 3-5: Form C1	47
Table A.5. Grades 3-5: Form C2	49
Table A.6. Grades 6-8: Form D1	51
Table A.7. Grades 6-8: Form D2	53
Table A.8. Grades 9-12: Form E1	55
Table A.9. Grades 9-12: Form E2	57
B. 2008-2009 Score Reports Interpretation Guide	59

List of Tables

		Page
Table 1.	Structure and Content of MontCAS ELP 2008-2009 Test Forms	4
Table 2a.	Configuration of MontCAS ELP 2006 and MontCAS ELP 2008 Forms	6
Table 2b.	Percent of MontCAS ELP 2008 Items (Points) from MontCAS ELP 2007 Forms	6
Table 3.	Summary of Reader Reliability for MontCAS ELP Constructed-response Items	13
Table 4.	Summary of MontCAS ELP 2008-2009 Item Difficulty and Discrimination by Grade span and Language Domain	17
Table 5.	Anchor Item Points by Form and Modality	19
Table 6.	Reliability, Raw Score and Scaled Score Descriptive Statistics for MontCAS ELP 2008-2009 Test Forms by Grade	24
Table 7.	Correlations Among Scaled Scores on Individual Language Domain Tests	29

List of Tables (Continued)

		Page
Table 8.	Numbers of Students Scoring at Each Performance Level on the MontCAS ELP test and on MontCAS Criterion-referenced Tests in ELA and Math	30
Table 9.	Correlations between Performance Levels on the MontCAS ELP and ELA and Math CRT	31
Table 10.	Performance on 2007 and 2008 MontCAS ELP Test Forms by Grade	32
Table 11.	Total MontCAS ELP Level by Grade in 2007 and 2008	36
Table 12.	Summary of 2007 to 2008 Growth Reports	38
Table 13.	Summary of 2006 to 2007 Growth Reports	39

List of Figures

		Page
Figures 1-5.	Step values of Anchor Items for 2007 and 2008 IELA forms	20

MontCAS English Language Proficiency (ELP) 2008-2009 Technical Report

1. Purpose of the Technical Report

The purpose of this report is to provide the Montana Office of Public Instruction (OPI) as well as Montana educators, citizens, researchers, and other interested parties with technical documentation for the development, administration, and reporting of the Fall 2008 administration of the MontCAS English Language Proficiency Assessment (MontCAS ELP). This report includes evidence of the reliability and validity of the assessment as well as other information about test administration and results. Although this technical report covers the 2008-2009 administration of the MontCAS ELP, some data from the previous administrations are included for reference and comparison.

2. Description of the MontCAS ELP

2.1 Purpose of the MontCAS ELP. The Montana English Language Proficiency Assessment (MontCAS ELP) is an assessment of English language proficiency for grades K-12. It is a modified version of an assessment developed for the Mountain West Consortium and designed to fulfill the requirements of "No Child Left Behind" (NCLB) legislation. The MontCAS ELP assesses English proficiency in Listening, Speaking, Reading, and Writing, and reports scores in each of those language domains as well as in Comprehension (a combination of select items from the Listening and Reading tests) and a total score, representing overall English proficiency. The MontCAS ELP was designed to assess the status of a student's proficiency in English and to measure progress in attaining English proficiency.

The MontCAS ELP was designed to be administered to all students who have been identified as "limited English proficient" (LEP) in the State of Montana. The process for identifying students as LEP is controlled at the district level and may include administering the Home Language Survey as well as one or more of a number of assessments. The instructions printed in the MontCAS ELP Test Administrator Manuals read as follows:

> "Montana observes the federal definition of limited English proficiency. Both language impact and academic achievement must be considered when identifying LEP students. A student must be identified as one of the following:
>
> 1. an individual who was not born in the U.S. or whose native language is a language other than English;
> 2. an individual who comes from an environment where a language other than English is dominant;

3. an individual who is American Indian or Alaskan Native and who comes from an environment where a language other than English has had a significant impact on the individual's level of English language proficiency.

The student must also have sufficient difficulty speaking, reading, writing, or understanding the English language to deny such an individual the opportunity to learn successfully in classrooms where the language of instruction is English or to participate fully in our society."

The LEP population in the state of Montana is different from that of many other states. In Montana, up to 80% of the students identified as LEP are of American Indian descent and are very likely growing up in a community where English is the primary language. The English used in that community may very well be a nonstandard version. The uniqueness of student populations in the Western United States, including the prevalence of students of American Indian descent, was part of the impetus for the formation of the Mountain West Consortium. And the test development procedures (Matthews, 2007) took the characteristics of the student population in member states into consideration. Although the population in Montana includes a higher percentage of students of American Indian descent, that population is not qualitatively different from that of other Mountain West member states.

2.2 Past and Present MontCAS ELP Forms. The first set of MontCAS ELP forms, designated MontCAS ELP Fall 2006, was administered in Fall 2006. These forms were based on Mountain West Form I and were previously administered in Idaho as the Idaho English Language Proficiency Assessment (IELA). More detailed information about these forms is included in the *MontCAS ELP Technical Report, 2006-2007*.

A second set of MontCAS ELP forms, designated MontCAS ELP 2007-2008, was administered in Fall 2007. The MontCAS ELP 2007-2008 forms were similar in structure to the MontCAS ELP 2006 forms but contained approximately 70% different items. The new items on MontCAS ELP 2007-2008 were developed as part of the original Mountain West Consortium item development and were drawn from the Mountain West item bank (i.e., Forms II and III). Item development was conducted in accordance with procedures outlined in Matthews (2007). New items were reviewed for content and structure and edited where appropriate. Directions for administration were revised, where necessary and appropriate, to conform to the conventions adopted in MontCAS ELP 2006. The MontCAS ELP 2007-2008 forms were previously administered in Idaho in Spring 2007 as the IELA. All edits to items were made in advance of the administration of the test in Idaho. Items that were in common between the 2006-2007 and 2007-2008 forms served as anchor items to equate the 2007-2008 to the 2006-2007 forms. More detailed information about these forms is included in the *MontCAS ELP Technical Report, 2007-2008*.

The structure of the 2008 MontCAS ELP forms, designated MontCAS ELP 2008-2009, differ somewhat from the structure of the 2006 and 2007 MontCAS ELP forms. The differences will be addressed in the next section. The forms administered as the MontCAS ELP 2008-2009 were previously administered in Idaho in Spring 2008 as the IELA. The items on these forms were drawn from those administered in 2006 and 2007. The version of the forms administered in Idaho included both operational and field test items. Prior to administration as the MontCAS ELP, all field test items were removed. Operationally, the MontCAS ELP 2008-2009 forms were identical to those administered in Idaho. Items that were in common between the 2007-2008 and the 2008-2009 forms served as anchor items to equate the 2008-2009 forms to the 2007-2008. More details of the equating are provided in a later section of this report.

2.3 Structure of the MontCAS ELP. MontCAS ELP test forms were designed for specific grade/grade clusters, K, 1-2, 3-5, 6-8, and 9-12, as shown in Table 1 (page 4). For every grade cluster except Kindergarten, there are two forms differentiated by a number suffix (e.g., C1 and C2). The Level 1 forms were designed to be administered to students on the lower end of the English proficiency scale (i.e., Beginner) and the Level 2 forms designed for students on the upper end of the scale (i.e., Intermediate and Advanced).

Table 1 shows for each test form the grade cluster in which it is administered and the numbers of items by item type in each language domain as well as the number of points represented by those items. The items and points in the Comprehension column do not contribute to the Totals shown in the last two columns because all Comprehension items are part of the Listening or Reading tests.

All Listening and Reading items were eligible to be included on the Comprehension test. Those items that assessed a lower-level reading skill (e.g., letter identification, sound-symbol correspondence) were not included as comprehension. In addition, stand-alone vocabulary items were not included although vocabulary-in-context items were included. Two individuals with extensive experience in test development independently identified those items on the Listening and Reading subtests that assessed comprehension. On those occasions where they disagreed, a third person evaluated the item to break the tie.

Table 1. Structure and Content of MontCAS ELP 2008-2009 Test Forms

Form	Grade Cluster	Item Type	Listen		Speak		Read		Write		Comp		Total	
			Itm	Pts	Itm	Pts	Itm	Pts	Itm	Pts	Itm	Pts	Itm	Pts
A	K	MC	3	3	-	-	10	10	-	-	3	3	13	13
		SA	12	12	7	7	17	17	-	-	15	15	36	36
		ER	-	-	3	8	-	-	-	-	-	-	3	8
		Total	**15**	**15**	**10**	**15**	**27**	**27**	**22***	**22***	**18**	**18**	**74**	**79**
B1	1-2	MC	15	15	-	-	15	15	-	-	23	23	30	30
		SA	-	-	7	7	-	-	11	11	-	-	18	18
		ER	-	-	3	8	-	-	2	4	-	-	5	12
		Total	**15**	**15**	**10**	**15**	**15**	**15**	**13**	**15**	**23**	**23**	**53**	**60**
B2		MC	18	18	-	-	18	18	-	-	35	35	36	36
		SA	-	-	6	6	-	-	8	8	-	-	14	14
		ER	-	-	4	12	-	-	3	10	-	-	7	22
		Total	**18**	**18**	**10**	**18**	**18**	**18**	**11**	**18**	**35**	**36**	**57**	**72**
C1	3-5	MC	15	15	-	-	15	15	4	4	27	27	34	34
		SA	-	-	7	7	-	-	5	5	-	-	12	12
		ER	-	-	3	8	-	-	2	6	-	-	5	14
		Total	**15**	**15**	**10**	**15**	**15**	**15**	**11**	**15**	**27**	**27**	**51**	**60**
C2		MC	18	18	-	-	16	16	8	8	34	34	42	42
		SA	-	-	6	6	1	2	-	-	1	2	7	8
		ER	-	-	4	12	-	-	3	10	-	-	7	22
		Total	**18**	**18**	**10**	**18**	**17**	**18**	**11**	**18**	**35**	**36**	**56**	**72**
D1	6-8	MC	15	15	-	-	15	15	6	6	29	29	36	36
		SA	-	-	9	9	-	-	3	3	-	-	12	12
		ER	-	-	2	6	-	-	2	6	-	-	4	12
		Total	**15**	**15**	**11**	**15**	**15**	**15**	**11**	**15**	**29**	**29**	**52**	**60**
D2		MC	18	18	-	-	14	14	10	10	32	32	42	42
		SA	-	-	6	6	-	-	-	-	-	-	6	6
		ER	-	-	4	12	2	6	3	10	2	6	9	28
		Total	**18**	**18**	**10**	**18**	**16**	**20**	**13**	**20**	**34**	**38**	**57**	**76**

Table 1. Structure and Content of MontCAS ELP 2008-2009 Test Forms (Continued)

Form	Grade Cluster	Item Type	Listen		Speak		Read		Write		Comp		Total	
			Itm	Pts	Itm	Pts	Itm	Pts	Itm	Pts	Itm	Pts	Itm	Pts
E1	9-12	MC	15	15	-	-	15	15	7	7	28	28	37	37
		SA	-	-	7	7	-	-	2	2	-	-	9	9
		ER	-	-	3	8	-	-	2	6	-	-	5	14
		Total	**15**	**15**	**10**	**15**	**15**	**15**	**11**	**15**	**28**	**28**	**51**	**60**
E2		MC	18	18	-	-	18	18	10	10	36	36	46	46
		SA	-	-	6	6	-	-	-	-	-	-	6	6
		ER	-	-	4	12	1	2	3	10	1	2	8	24
		Total	**18**	**18**	**10**	**18**	**19**	**20**	**13**	**20**	**37**	**38**	**60**	**76**

* Items on the Kindergarten Writing test are configured as a checklist completed by the test administrator.
MC - Multiple Choice; SA - Short Answer; ER - Extended Response

Table 2a (page 6) compares the structure of MontCAS ELP 2008 forms to those administered in 2006 (which were structured identically to MontCAS ELP 2007). This table clarifies some of the differences between MontCAS ELP 2008 forms and those administered in 2006 and 2007. First, MontCAS ELP 2008 forms were shorter in terms of number of points per language domain than their predecessors. This shortening was related to several of the other changes detailed as follows. Second, whereas in previous versions of MontCAS ELP, the same Speaking and Listening items appeared on Level 1 and Level 2 forms within a grade cluster, on MontCAS ELP 2008, the majority of items on Level 1 Speaking and Listening tests within each grade cluster were different from those on the Level 2 Listening and Speaking tests (i.e., only Level 1 to Level 2 linking items were common). Third, the difficulty of the MontCAS ELP 2008 forms was adjusted to align Level 2 forms more closely with the abilities of students to whom they were being administered. Results of both MontCAS ELP 2006 and 2007 suggested that those forms were not challenging enough to capture performance at the upper levels of English language proficiency.

Table 2a. Configuration of MontCAS ELP 2006 and MontCAS ELP 2008 Forms

Year	Form	L		S		R		W		C		Total	
		Itms	Pts	Itms	Pts	Itms	Pts	Itms	Pts	Itms	Pts	Itms	Pts
2006	**B1**	22	22	14	22	15	15	13	15	31	31	64	74
	B2	22	22	14	22	20	20	13	20	39	39	69	84
2008	**B1**	15	15	10	15	15	15	13	15	23	23	53	60
	B2	18	18	10	18	18	18	11	18	35	35	57	72
2006	**C1**	22	22	14	22	15	15	11	15	31	31	62	74
	C2	22	22	14	22	19	20	12	19	38	39	67	83
2008	**C1**	15	15	10	15	15	15	11	15	27	27	51	60
	C2	18	18	10	18	17	18	11	18	35	36	56	72
2006	**D1**	22	22	14	22	15	15	11	15	32	32	62	74
	D2	22	22	14	22	20	24	13	20	40	44	69	88
2008	**D1**	15	15	11	15	15	15	11	15	29	29	52	60
	D2	18	18	10	18	16	20	13	20	34	38	57	76
2006	**E1**	22	22	14	22	15	15	11	15	32	32	62	74
	E2	22	22	14	22	21	25	13	20	41	45	70	89
2008	**E1**	15	15	10	15	15	15	11	15	28	28	51	60
	E2	18	18	10	18	19	20	13	20	37	38	60	76

Table 2b shows by form and language domain the percent of MontCAS ELP 2008 items that appeared on MontCAS ELP 2007 forms. This table shows that there was considerable variability in the percent of items drawn from the same test at the same level in the prior year.

Table 2b. Percent of MontCAS ELP 2008 Items (Points) from MontCAS ELP 2007 Forms

Form	L	S	R	W
A	100	100	100	100
B1	87	73	100	60
B2	72	50	72	61
C1	33	93	100	53
C2	72	50	78	67
D1	40	67	67	80
D2	72	61	75	95
E1	47	47	73	33
E2	72	80	60	65

2.4 Alignment of the MontCAS ELP. An alignment study of the MontCAS ELP to the Montana English Language Proficiency Standards has not yet been completed. In the development of the Mountain West Consortium Test (Matthews, 2007), the member states of the consortium developed a set of common English language development (ELD) standards. The MWAC ELD standards were used to guide item development for the Mountain West Test.

3. MontCAS ELP 2008-2009 Administration

3.1 Testing Window. The testing window for MontCAS ELP 2008-2009 was October 20 through November 21, 2008. All test materials were to be returned to Questar by December 12, 2008.

3.2 Assessment Training. To prepare systems for the administration of the Fall 2008 MontCAS ELP, a *Training PowerPoint® Presentation* was created to cover three main areas: What's New, Test Administration, and Post-Test Instructions. A Training CD with this presentation was shipped to all systems with a known LEP population on September 4, 2008 and a PDF version of the presentation (showing each slide and the notes section) was posted to the Office of Public Instruction website, http://www.opi.mt.gov/curriculum/MontCAS/#gpm1_11. A *Training Handout*, which showed each slide from the Training Presentation, was also provided. The General Instructions from each *Test Administrator Manual* (Form A, Form B, Form C, Form D, and Form E) were also posted on the OPI website to allow test coordinators a chance to begin preparing before assessment materials arrived.

Each System Test Coordinator was encouraged to read through these presentations prior to administration and to consider using the PowerPoint® presentation to train test administrators.

To prepare for testing, test administrators were instructed (in the *Test Administrator Manual*) to:

- read the manual completely;
- ensure that they had adequate materials for all students who would be tested;
- notify students in advance of testing;
- print students' first and last names on the answer document; and
- secure a CD player (or computer with CD-ROM drive, sound card and speakers) for administering the Listening test, and check the sound quality.

3.3 Test Administrator Scripts. Specific step-by-step instructions and script were provided for each test form in a *Test Administrator Manual* specific to that particular form. Scoring guides were provided for all oral constructed responses. Such items occurred throughout the Kindergarten form, but only in the Speaking test at all other grade spans. Where appropriate, examples of full-credit and partial-credit responses were provided.

3.4 Listening Test Administration. The Listening test was administered with a CD recording. This ensured that all students heard the questions in the same voice and at the same pace. The recording included a tone after each question signaling the test administrator to pause the CD while students responded. A printed Listening Script for each form was available to any school that requested it.

3.5 Setting for the Test. For the individually administered subtests, test administrators were advised as follows: "The test setting should be a quiet one-to-one environment. The testing should take place where other students cannot hear or see the testing materials. The test administrator should sit close enough to the student to point to questions and illustrations in the student's test booklet during test administration."

For the group-administered subtests, test administrators were advised as follows: "The test setting for the group-administered sections is a quiet classroom. The students should have in front of them only their test booklet, answer document, and a No. 2 pencil."

3.6 Timing. The MontCAS ELP is an untimed test and test administrators were advised to allow students as much time as they needed to finish any given subtest.

3.7 Prompting and Repeating Test Information. The following rules regarding prompting or repeating information were printed in all *Test Administrator Manuals*:

Prompting is the provision of additional information to students during administration of the assessment. Prompting includes

- elaborating on questions,
- clarifying information provided in reading selections or any test question,
- pointing out specific information in the questions or graphics,
- providing cues that might normally be part of an instructional strategy, and/or
- suggesting strategies that a student may use to arrive at a correct response.

In general, prompting is **not** allowed in this test because it may give an unfair advantage to some students. However, in specific situations where partial or unclear responses are given, the following general prompts are appropriate.

To clarify the student's response, the test administrator may say,
I don't understand what you said.
Can you tell me more?

If the student answers in another language, the test administrator may say,
Can you say that in English?

The test administrator may repeat directions, if necessary, but must do so before the child begins a response.

If there is a distraction or interruption, the selection or question may be repeated.

If a student asks for a question to be repeated, the test administrator may repeat the question only once.

If the student still does not understand what is being asked, the test administrator should score that question as though the student gave no response (*BL*).

The test administrator must not modify directions in any way. To do so would provide an unfair advantage to one student or a group of students over others.

The test administrator should allow approximately 15 seconds of wait time for a student to begin a response to a question. This gives the student time to gather his or her thoughts and to think carefully before responding in English. If a student has not responded after 15 seconds, the test administrator should move on to the next item or task and score the item as "no response" (*BL*).

3.8 Testing Absentees. Test administrators were advised to make every effort to see that all LEP students in the school were administered all sections of the MontCAS ELP. If a student was absent for a particular testing session, a make-up test was to be scheduled, as long as it was within the testing window.

3.9 Testing Accommodations. For visually impaired students, the MontCAS ELP 2008-2009 was available (by special order) in Braille and in Large-Print. No Braille forms were ordered before the August 22, 2008 deadline. There were two orders for the C2 Large Print test booklets and one order for the D2 test booklet.

Detailed guidelines for Standard and Nonstandard Accommodations were provided in each *Test Administrator Manual*. In the Guidelines for Standard and Nonstandard Test Accommodations section, it was noted that some of the accommodations were crossed out on the listing and NA was coded in the accommodations section of answer documents. These crossed-out accommodations were not appropriate for MontCAS ELP students. The guidelines included the statement:

> "The fact that the MontCAS ELP is an untimed test and that there is considerable graphic support should help with increased comprehension for LEP students, including LEP

> students with special needs. However, in some cases it may be necessary to provide specific accommodations."

Test administrators were instructed to only bubble accommodations IF the accommodation was made for a student with special needs. Standard accommodations for the MontCAS ELP were available to students with an Individual Education Plan (IEP) or 504 Plans and to "all students if the accommodation(s) had been part of the student's classroom routine three months prior to testing." Standard accommodations were to be "determined on an individual basis, student by student, rather than for groups of students." Nonstandard accommodations were available only for a student with IEP/504 plans and caution was advised for teams in considering whether a student required a nonstandard accommodation.

Test administrators were warned that such accommodations should be used only when absolutely necessary. If a student was tested with accommodations, the test administrator was instructed to mark the appropriate bubble (Box 7) on the answer sheet.

Certain accommodations would necessarily invalidate test scores. The following list of non-allowable accommodations was provided in the *Training PowerPoint® Presentation* and the *Training Handout*:

> The following accommodations are NOT allowed:
> - Test administration in a language other than English.
> - Translation of the assessment into another language.
> - Translation of the assessment into sign language.
> - Use of dictionaries or other reference aids. This includes both monolingual and bilingual dictionaries.
> - Accepting responses in a language other than English.
>
> (If students respond in their native language, the test administrator may ask them if they can "say that in English." If they cannot, the response counts as 0.)
> The use of any of the non-allowable accommodations will invalidate test scores.

4. MontCAS ELP 2008-2009 Test Security

System Test Coordinators were instructed to "communicate the importance of maintaining test security" before, between, and after testing sessions. Additionally, in the *Test Administrator Manuals*, it stated that "No part of any test booklet or test administrator manual (except where expressly stated) may be reproduced or transmitted in any fashion. At the conclusion of the test

administration, all test materials (both used and unused) must be accounted for and returned to the System Test Coordinator, who will return all materials to Questar Assessment, Inc."

4.1 Bar-Coding and Return of Secure Materials. All secure materials (test booklets, prompt books, Listening test CDs, and test administrator manuals) were individually bar-coded. These secure test materials were scanned upon packing and distributing to systems and then scanned again upon return to Questar to account for materials. Test Coordinators were instructed to return all test materials—used and unused—to Questar. A detailed description of the check-in of secure materials is included in the *2006-2007 Technical Report*.

4.2 Storage and Shredding of Secure Materials. After scoring, all used test booklets and answer documents were stored in Questar's secure warehouse facility in Apple Valley, Minnesota. Used answer documents are stored according to their processing for quick retrieval, if necessary. Access to these facilities is limited to Questar staff. Used student answer documents and unused and non-scannable secure materials must be stored for 180 days, and then Questar requests written permission from the State Manager to recycle the materials using a secure method of destruction. Questar received written permission from the Montana Office of Instruction in September 2009 to destroy the 2008-2009 materials, except for file copies.

5. MontCAS ELP 2008-2009 Scoring and Reporting

5.1 Scoring of Multiple-Choice Items. Multiple choice items (which are bubbled on the student test booklet or answer document) were scored electronically. One (1) point was given for the correct answer bubbled. Zero (0) points were given for incorrect answer bubbled or multiple bubbles marked. If no item was bubbled (an omit), the response was scored as a "blank."

5.2 Writing Checklist. The Writing raw score for (Kindergarten level) Form A was calculated as follows: 1 point was allocated for each skill on the Writing Checklist that the student "does most of the time" or of which they "demonstrate mastery." Thus, the Writing Checklist generated a maximum raw score of 22 points.

5.3 Scoring of Constructed-Response Items. The MontCAS ELP includes constructed-response (CR) items (separated into short answer [SA] and extended response [ER] in Table 1, page 4) in Speaking and Writing as well as a few CR items in Reading. Speaking CR items were scored by the test administrator at the time of test administration. Scoring guides and examples of full and partial-credit items were included as part of the *Test Administrator Manual*. Speaking responses were not recorded and no attempts were made to assess the validity or reliability of the rating of Speaking items.

Writing and Reading constructed-response items were scored at the Questar scoring center using a 1-point, 2-point, or 4-point scale. The table that follows shows the grade spans, forms, levels, and Domains where there are constructed-response items. A second independent read was provided for 20% of the Level 2 constructed response items. Level 1 constructed-response items were rated by the Questar Scoring Directors without a rescore due to the low quantities and due to the non-scannable test booklets/answer documents used for Level 1 forms.

CONSTRUCTED-RESPONSE ITEMS		
Grade Span	**Forms**	**Level and Domain**
1-2	Form B	Level 1 & Level 2 Writing
3-5, 6-8, 9-12	Forms C, D, E	Level 1 & Level 2 Writing; Level 2 Reading

Training Materials. A *Scoring Manual for Open-Ended Reading/Writing Responses* was used in the training of readers for scoring constructed-response items. A separate scoring manual was created for each grade span (B, C, D, and E). Questar's content specialists reviewed the scoring guides and rubrics for the constructed-response items, noted where there were weaknesses (if any) in the rubrics, and identified types of responses that will likely be seen in the operational responses. When necessary, sample responses were added to various items and score points to present a more complete scoring guide (which consists of background information, the scoring rubrics, and annotated anchor responses) used to train readers.

Staffing. The scoring team consisted of one scoring director and 10 readers. The director managed scoring of reading and writing items. All readers scored the entire project (reading and writing items). None of the readers was released during training or subsequent scoring due to poor performance. Readers were trained on each item by grade span prior to scoring any of the items in that grade span. Following the group training, the readers completed paired reads on individual items. As the scoring proceeded, Reader Reliability Statistics and Scorepoint Distribution Statistics were monitored for each reader on a daily basis.

Reader Reliability. The constructed-response items that were scored by two readers provide information on reader reliability. Data relevant to this issue are summarized in Table 3 (page 13). This table shows, for each Level 2 form for each item or set of items, the maximum point value of the item(s) (Pts), the number of student papers read twice (N), the percent of items on which the readers agreed exactly (% Exact), and the percent of items on which reader agreement was

within +/1 one score point (% Ex+Adj). All items, even those with maximum point values of 4, were at or near 100% exact + adjacent agreement.

Table 3. Summary of Reader Reliability for MontCAS ELP Constructed-response Items

Form	Domain	Item(s)	Pts	N	% Exact	% Ex + Adj
B2	W	1-2	1	356	98	100
		3-4	1	356	96	100
		5	1	356	95	100
		6	1	356	93	100
		7	1	356	100	100
		8	1	356	99	100
		9	2	356	88	100
		10	4	356	93	100
		11	4	356	90	99
C2	W	9	2	556	82	100
		10	4	556	89	99
		11	4	556	79	99
	R	17	2	556	91	99
D2	W	11	2	408	85	100
		12	4	408	88	100
		13	4	408	83	99
	R	11	2	408	85	100
		16	4	408	86	100
E2	W	11	2	342	88	100
		12	4	342	81	99
		13	4	342	87	100
	R	15	2	342	78	99

Handscoring Issues. Students writing constructed responses outside of a designated response area is a recurring issue on the MontCAS ELP. This issue arose once again during the scoring of constructed response items on the 2008-2009 MontCAS ELP. These item mismatches were observed on both 2006-2007 and 2007-2008 administrations of the MontCAS ELP. In each subsequent administration, the issue was addressed in the training for test administrators. In December 2008, OPI and Questar reviewed the issue of item mismatch in the context of a discussion of the constructed response scoring process, and agreed on rules for handling the handscoring of these occurrences of mismatch during the 2008-2009 administration. The rules are outlined below.

- If there was demonstrated intent on the answer document to indicate that a response had been misplaced, handscoring would score the response for the misplaced item.

Demonstrated intent was defined as instances where (1) the student/teacher wrote in the item number next to the written response, or (2) student/teacher noted that the item(s) was written in the wrong spot(s).

- If there was no demonstrated intent, the misplaced response was considered off-topic and would be scored "0" points.
- If there were two responses in the designated area for a single item and no demonstrated intent, the response that matched the prompt would be scored. If the next response was blank and no intent was communicated, that item was scored as a BL (blank).

During handscoring, there were five instances of item mismatches where the teacher or student included a note alerting the reader to the mismatch. Questar readers were then able to apply ScorePoint's item mismatch feature and score the response for the misplaced item. There were other instances of item mismatch where there was no demonstrated intent by the student or test administrator. Those instances were handled in the following ways in accordance with the rules outlined above:

- **Two responses written on one page (with no note alerting to a mismatch).** Readers scored the response that matched the item corresponding to that page and did not count the "extra" writing against the student's score.
- **Items written completely on wrong page (with no note alerting to a mismatch).** Readers scored a "0" for off topic if it was off topic.
- **Blank pages.** Readers scored a "B" for Blank.

For the 2009-2010 administration, a decision was made to change some design elements on the answer documents in an attempt to make it more clear where students should record their responses. These changes include:

- Printing item numbers in boldface within the top-left of the response box.
- Providing additional space between response boxes, if possible.
- At top of response box, adding the message "Answer question <number> here."
- At bottom of page, where applicable, changing "Turn page to continue" to "Turn page to answer next question".
- At the beginning of the instructions in the *Test Administrator Manual*, adding the message, "Write only one response in each response box."

5.4 Reporting. Student performance in each of the language domains (Listening, Speaking, Reading, Writing) and Comprehension was reported in terms of raw score, scaled score, and proficiency levels. Student performance was also reported on the overall (Total MontCAS ELP) test in terms of raw score, scaled score, and proficiency level. In February 2007, a panel of

Montana educators met to set standards for the MontCAS ELP in the form of cut scores for each proficiency level by grade. A detailed description of standard setting procedures is included in the *2006-2007 Technical Report*. The reported scores were defined in the *2008-2009 MontCAS ELP Assessment Score Reports Interpretation Guide*. A copy of that guide is included as **Appendix B**.

Incomplete Testing. Students were required to take all four language domain tests. If a student did not take one or more of the domain tests, the reports showed dashes in place of scores for that domain. The reported Total MontCAS ELP score was based on the domain tests for which there are scores. Thus, if a student failed to take the Speaking Test for whatever reason, the Total MontCAS ELP score was based on a raw score of zero in Speaking. The reported Comprehension scores–which were based on a subset of Listening and Reading scores–was affected in the same way if the student failed to take either the Listening or Reading Test.

Reports Shipment. MontCAS ELP 2008-2009 results packages were shipped to systems on February 20, 2009. The system and each of its schools had separate results packets. The reports that were in each packet are listed below. In addition, copies (1 copy for each school and system) of the *2008-2009 MontCAS ELP Assessment Score Reports Interpretation Guide* (SRIG) were included in the shipment. The SRIG included a sample of each report type with information for understanding the report and information for using the MontCAS ELP results. The SRIG was also posted on the OPI website, http://www.opi.mt.gov/curriculum/MontCAS/#gpm1_11.

MontCAS ELP System Packet – 2008-2009

- Contents Sheet
- System Summary Reports by grade
- System Growth Reports by grade
- Copy of each School Summary Report
- Copy of each School Roster

MontCAS ELP School – 2008-2009

- Contents Sheet
- School Summary Reports by grade
- School Rosters
- Individual Student Reports
- Student Labels
- Parent Reports

Note that the System Growth Report showed growth within the system for those students who were assessed with the MontCAS ELP in 2006-2007, 2007-2008, and 2008-2009 (confirmed by a State ID # match). Growth Reports were provided only when there were 10 or more students

per report. If there were fewer than 10 students, system personnel were instructed to examine the student's Individual Student Reports to determine growth.

6. MontCAS ELP 2008-2009 Item Analyses

This section provides classical item-level statistics for all items administered on MontCAS ELP 2008-2009 forms. The p-value is presented as an index of item difficulty and the point-biserial correlation is presented as an index of item discrimination.

P-Values. For multiple-choice items, the p-value statistic is defined as the proportion of students that answer an item correctly. For constructed-response items, the p-value is reported as the average number of points out of the maximum number of possible points for an item. P-values range from zero to one (1.0). A high p-value means that an item is easy; a low p-value means that an item is difficult. Generally, it is desirable for tests to include items that span a range of difficulty.

Point-biserial correlations. The point-biserial correlation for each item is an index of the association between the item score and the total-test score. It shows how well the item discriminates between low-ability and high-ability students, where ability is inferred from the overall test score. Point-biserial correlation coefficients range between −1.0 and +1.0. High positive values indicate that a high-ability student is more likely (than a student with lower ability) to answer an item correctly and low negative values indicate that a low-ability student is more likely (than a student with higher ability) to answer an item correctly.

Table 4 (page 17) shows the average p-value and range and median point-biserial correlation coefficients and range by language domain and test form. These data are only shown for Level 2 forms because the numbers of Level 1 forms administered were low even when aggregated across grades within a grade span. As in previous administrations of the MontCAS ELP, there were differences in both range and average p-values across language domains. The average p-values in Reading and Writing remain lower than the averages in Listening and Speaking, although the differences are smaller than in either 2006-2007 or 2007-2008. The difference in the current administration is most likely due to the reconfiguration of Listening and Speaking tests.

Table 4. Summary of MontCAS ELP 2008-2009 Item Difficulty and Discrimination by Grade span and Language Domain

Grade Span	Form	Domain	N	Item p-value		Point Biserial	
				Avg	Range	Med	Range
K	A	L	402	0.63	0.36 - 0.88	0.37	0.22 - 0.48
		S	402	0.75	0.40 - 0.92	0.35	0.25 - 0.53
		R	402	0.43	0.19 - 0.77	0.45	0.27 - 0.66
		W	402	0.27	0.04 - 0.74	0.45	0.18 - 0.57
1-2	B2	L	854	0.66	0.37 - 0.92	0.29	0.14 – 0.42
		S	854	.83	0.63 - 0.95	0.39	0.30- 0.54
		R	854	0.65	0.41 - 0.87	0.35	0.20 - 0.47
		W	854	0.47	0.21 - 0.69	0.45	0.24 - 0.63
3-5	C2	L	1,215	0.67	0.28 - 0.92	0.32	0.10 - 0.39
		S	1,215	0.87	0.73 - 0.96	0.30	0.24 - 0.49
		R	1,215	0.64	0.41 - 0.86	0.40	0.19 - 0.43
		W	1,215	0.62	0.24 - 0.77	0.36	0.30 - 0.52
6-8	D2	L	936	0.73	0.47 - 0.94	0.32	0.18 - 0.43
		S	936	0.87	0.73 - 0.96	0.32	0.25 - 0.46
		R	936	0.56	0.33 - 0.82	0.33	0. 16 - 0.52
		W	936	0.69	0.42 - 0.94	0.36	0.19 - 0.58
9-12	E2	L	924	0.67	0.40 - 0.87	0.39	0.19 - 0.48
		S	924	0.83	0.68 - 0.94	0.38	0.24 - 0.46
		R	924	0.66	0.31 - 0.94	0.36	0.13 - 0.51
		W	924	0.66	0.33 - 0.87	0.37	0.23 - 0.56

Tables with item difficulty and discrimination data by item are included as **Appendix A** (page 41)**.** Analyses of test level data, including raw score descriptive statistics and test reliability measures, are reported in Table 6 (page 24).

7. Scaling and Equating of the MontCAS ELP

Initial scaling and equating of the 2008-2009 MontCAS ELP forms were completed on those forms when they were administered in Spring 2008 as the Idaho English Language Proficiency Assessment. The decision was made in 2006 to use the Idaho data for item calibration, scaling and equating because the population to whom the forms were administered in Idaho was larger than the population to whom the test was administered in Montana. Although the LEP populations in Idaho and Montana are significantly different (approximately 85% of LEP students in Idaho are of Hispanic origin whereas approximately 85% of LEP students in Montana are of American Indian origin), concerns about the small size of the sample in Montana outweighed concerns about differences in the student population. A brief summary of the equating procedures follows.

The MontCAS ELP 2008 test forms, following their administration in Idaho as the IELA 2008, were equated to MontCAS ELP 2007 forms, administered as the IELA 2007, so that scores could be reported on the same score scale. Prior to equating 2008 to 2007 forms, however, 2008 items in each grade cluster test form were calibrated using the Rasch Partial Credit Model (PCM), as implemented in WINSTEPS, version 3.57.1. This model, which was used to calibrate 2007 items, is appropriate for short-answer and constructed response items on the Speaking and Writing subtests as well as multiple-choice items administered across the language domains. As a first step, items on 2008 forms A, B2, C2, D2, and E2 were calibrated, with items on each grade-cluster form calibrated independently. Items on 2008 Level 1 forms, B1, C1, D1, and E1, were then calibrated by fixing the item parameters for those items that are common between the two Levels of each grade cluster (i.e., forms C1 and C2) to the same values as the Level 2 calibration for those items. This calibration procedure equated Forms B1, C1, D1, and E1 to Forms B2, C2, D2, and E2, respectively, ensuring that, within each grade cluster, scores on the Level 1 and Level 2 forms are reported on the same scale.

Following the item calibration, MontCAS ELP 2008 test forms were equated to the 2007 forms using a common item or anchor test design. Anchor items, those items that appeared in identical format in both the Spring 2007 form and in the Spring 2008 form (as administered in Idaho as the IELA), were embedded in Forms A, B2, C2, D2, and E2. Within each grade cluster, at least 30% of the items were in common between the 2007 and 2008 forms. The numbers of common items by form and language domain are shown in Table 5 (page 19). For each language domain and the Total MontCAS, this table shows in the “B” column the number of points represented by items in common between the 2007 and 2008 forms. In the course of the equating, detailed in a subsequent section, some of the items represented in the “B” column were not used as linking items. Whereas all of the items remained on the test, only those shown in the “A” column were used as equating items. The procedure used to determine whether to use a common item in the equating is detailed in the following paragraph.

Table 5. Anchor Item Points by Form and Modality

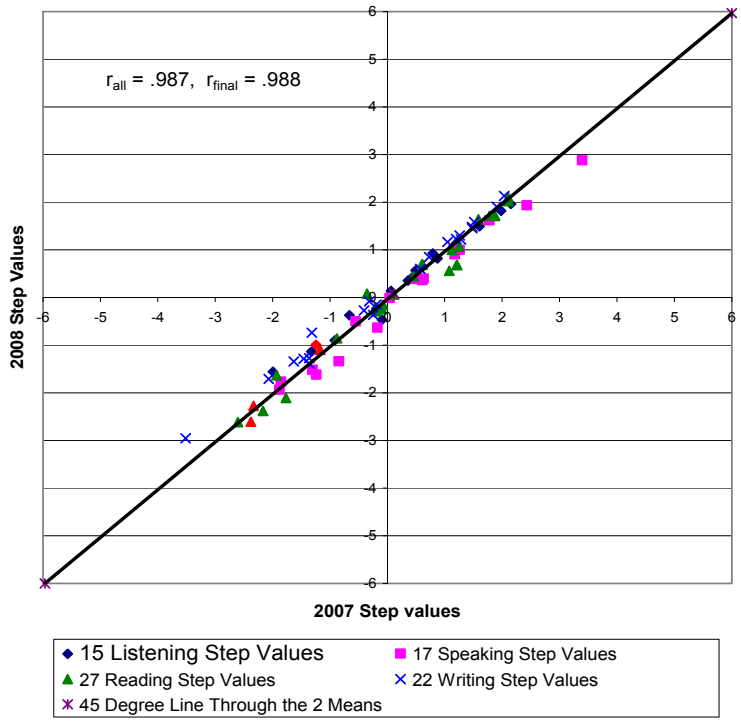
Form	Listening		Speaking		Reading		Writing		Total	
	B	A	B	A	B	A	B	A	B	A
A	17	17	15	15	27	27	22	20	81	79
B2	9	9	9	9	8	8	11	5	37	31
C2	10	10	10	8	13	13	10	6	43	37
D2	11	11	11	11	16	16	17	9	55	47
E2	11	11	11	11	11	11	13	7	46	40

Prior to equating 2008 to 2007 forms, each anchor item was evaluated for stability. As part of that evaluation, the calibrated difficulty (step value) of each anchor item in the current year (2008) was plotted against the calibrated difficulty of that item in the prior year (2007). Ideally, these plots should fall on a 45-degree line, indicating that calibrated values are stable from year to year. Those points that fall quite far from the line are referred to as outliers. For the anchor items in each of the five forms, the 2008 step values were plotted against the 2007 step values and these plots are shown in Figures 1-5. The numbers of plotted points for Forms A, B2, C2, D2, and E2 are 81, 37, 43, 55, and 46, respectively. Generally, the step values fell along this 45-degree line as the model requires. Of course, not all points are on or right next to the line due to error that is inherent in all measurement, and occasionally, a point is quite far from the line. Across the five forms, there were only a few outliers and these outliers were removed from the equating. Once the items were initially equated, a difference was calculated between the two step values (2007 step value – 2008 step value). Outliers were defined as items with an absolute difference of 0.60 logits or greater. The items that were not included as part of the equating were still scored and were used as operational items on their respective forms, but those items were not included in the calculations to determine final equating constants. (Note that when a constructed response item with multiple score points had at least one outlier point, the entire item was removed from the equating.) After deleting items with outlier values, the number of step values for the forms as listed previously is 79, 31, 37, 47, and 40.

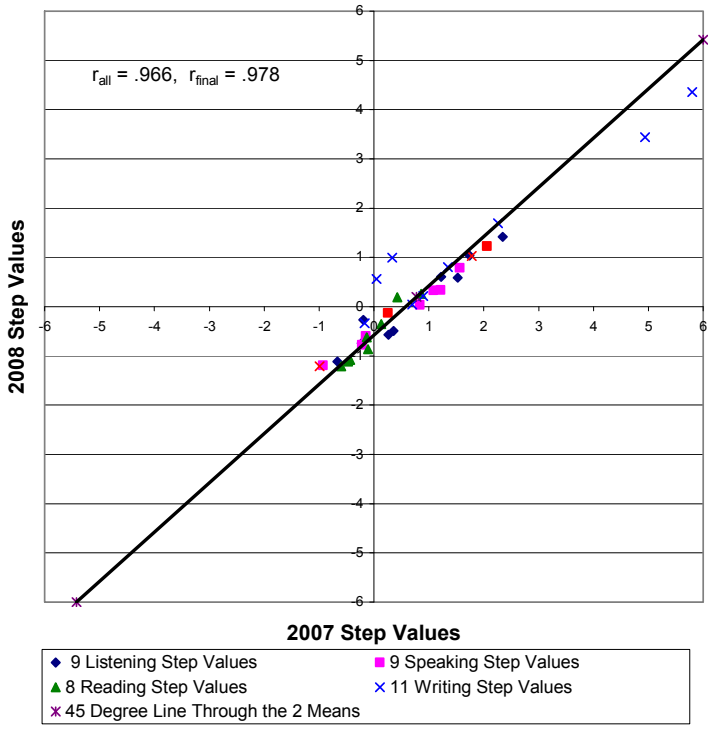
In Figures 1 through 5, two correlation coefficients (r) are given in the upper left-hand corner of each plot: one for all anchor items and the other for the final anchor items with outliers removed.

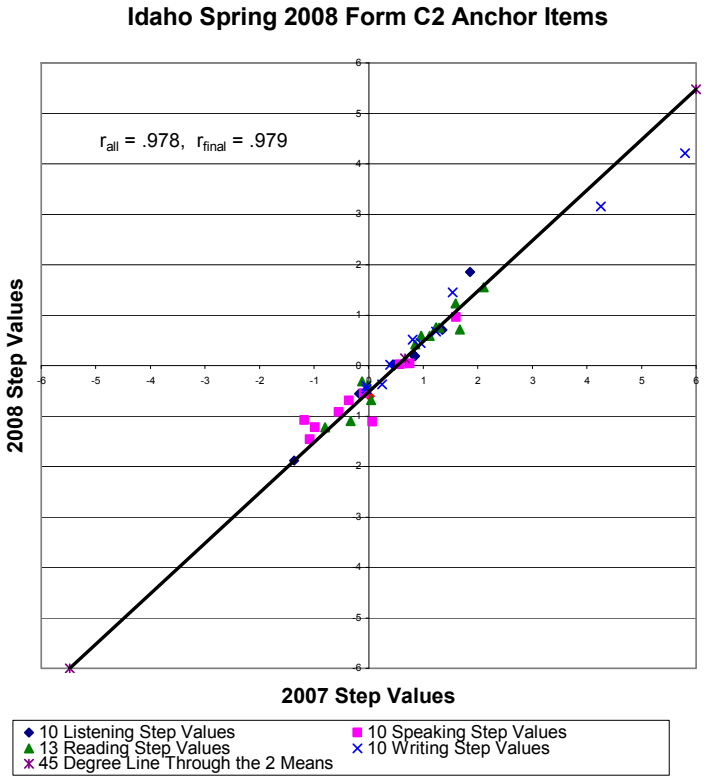
Figures 1-5. Step values of Anchor Items for 2007 and 2008 IELA forms.

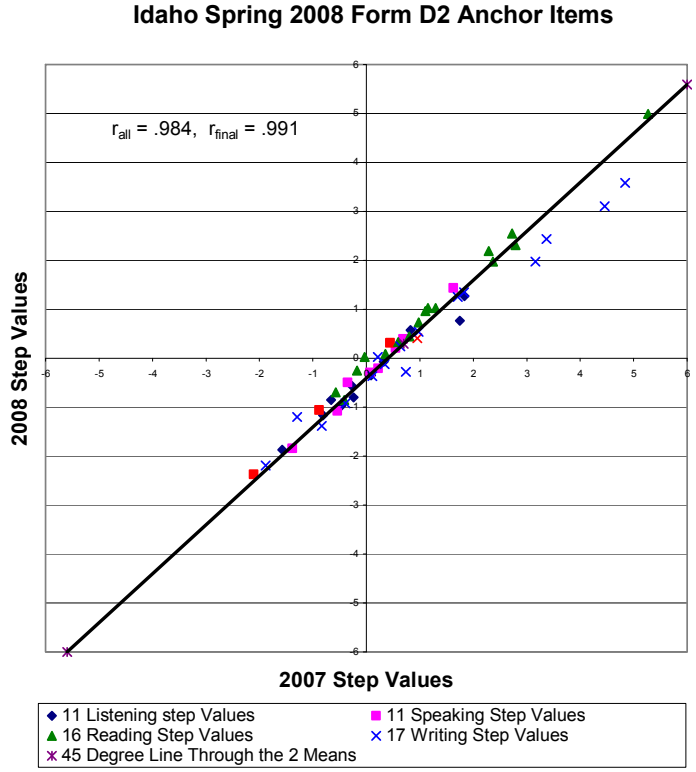
Idaho Spring 2008 Form A Anchor Items

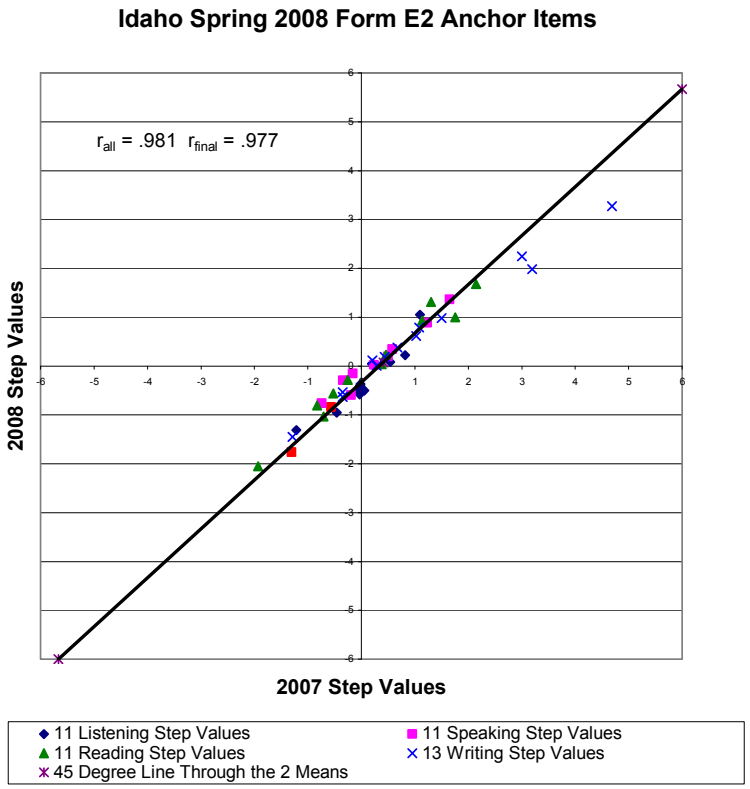


Idaho Spring 2008 Form B2 Anchor Items




Idaho Spring 2008 Form C2 Anchor Items
r_{all} = .978, r_{final} = .979
2008 Step Values
2007 Step Values
10 Listening Step Values
10 Speaking Step Values
13 Reading Step Values
10 Writing Step Values
45 Degree Line Through the 2 Means


Idaho Spring 2008 Form D2 Anchor Items
r_{all} = .984, r_{final} = .991
2008 Step Values
2007 Step Values
11 Listening step Values
11 Speaking Step Values
16 Reading Step Values
17 Writing Step Values
45 Degree Line Through the 2 Means



With the outliers removed, the final anchor items were used to develop a linking constant for each form that places the item step values from the 2008 form on the same Rasch logit scale as the 2007 form. The linking constant was computed as the difference between the average step value from the 2007 form's Winsteps calibration, minus the average step value from the 2008 form's Winsteps calibration. Adding this linking constant to the step values for each of the items in the 2008 form places all of the 2008 form's step values (and log ability estimates) on the same Rasch logit scale as the 2007 form. A separate linking constant was calculated for each grade cluster and applied. This constant was applied to items on all forms including both the Level 1 and Level 2 forms.

Once all items from the 2007 and 2008 forms were placed on the same logit difficulty scale established in 2006, scaled scores were computed for the 2008 forms. A linear transformation that was developed in the first year for each grade cluster form and test was applied to the equated Rasch log ability scale for the 2008 grade cluster form to yield equated scaled scores. The same cut scores for each proficiency level and grade established in the 2007 MontCAS ELP standards setting were applied.

8. Reliability of the MontCAS ELP

Data bearing on the reliability of MontCAS ELP 2008-2009 test forms are shown in the panels of Table 6 (page 24). This table shows for each form and each language domain (and Comprehension and the Total Test) the number of students (N) who were administered the form, coefficient Alpha, a measure of internal-consistency reliability, the maximum raw score attainable, and the mean, standard deviation, and standard error of measurement (SEM) in both raw score and scaled score units. Number of students represents the number for whom there was a valid test score and may vary across language domains in a grade to the extent that there were students who did not attempt one or more of the language domain tests. There is a total score for each student regardless of whether or not all language domain tests were attempted. Data are aggregated by grade for Level 2 forms but by grade span for Level 1 forms due to the small numbers of students administered the latter.

As reported for the 2006-2007 and the 2007-2008 administrations of the MontCAS ELP, there is some variability in the alphas over tests and grades. Low values (alpha < 0.70) do not seem to be restricted to a particular grade span, a particular test level (i.e., 1 or 2) or to a language domain. For the total test, the level at which classification decisions are made, reliability is consistently high (alpha >0.80) across forms and grade levels.

Table 6. Reliability, Raw Score and Scaled Score Descriptive Statistics for MontCAS ELP 2008-2009 Test Forms by Grade

Grade K				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
A	Listening	402	0.81	15	9.4	3.6	1.55	106.5	22.9	9.97
	Speaking	402	0.70	15	11.0	3.6	1.95	106.1	19.5	10.69
	Reading	402	0.91	27	11.7	6.9	2.06	83.7	28.8	8.60
	Writing	402	0.93	22	5.9	5.5	1.48	67.4	32.8	8.85
	Comprehen	402	0.81	18	9.8	3.9	1.71	104.5	20.0	8.70
	Total	402	0.94	79	38.0	15.2	3.75	383.0	31.8	7.84

Grades 1-2				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
B1	Listening	109	0.68	15	10.2	2.6	1.46	97.5	15.7	8.85
	Speaking	109	0.76	15	9.1	3.9	1.91	101.9	22.8	11.18
	Reading	110	0.65	15	8.9	2.8	1.64	86.7	15.8	9.41
	Writing	109	0.85	15	6.9	3.9	1.52	90.7	22.8	8.80
	Comprehen	110	0.70	23	14.0	3.8	2.05	91.7	13.1	7.17
	Total	110	0.88	60	34.9	10.5	3.56	383.3	37.0	12.61

Grade 1				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
B2	Listening	383	0.60	18	10.5	2.8	1.80	103.1	12.1	7.64
	Speaking	384	0.70	18	13.0	4.0	2.20	108.6	21.3	11.70
	Reading	383	0.65	18	9.4	3.1	1.84	97.1	12.5	7.44
	Writing	382	0.71	18	4.6	2.8	1.52	88.6	16.6	8.97
	Comprehen	385	0.74	36	19.3	5.1	2.60	99.7	10.6	5.46
	Total	386	0.83	72	37.2	9.7	3.95	398.5	27.9	11.42

Grade 2				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
B2	Listening	464	0.60	18	13.2	2.4	1.51	114.8	12.2	7.75
	Speaking	468	0.65	18	14.4	3.8	2.24	116.3	23.4	13.87
	Reading	468	0.65	18	13.5	2.7	1.63	115.4	15.0	8.88
	Writing	466	0.74	18	9.2	3.3	1.68	111.6	16.3	8.42
	Comprehen	468	0.74	36	26.0	4.5	2.31	114.2	11.7	6.00
	Total	468	0.82	72	50.1	9.3	3.91	435.7	29.3	12.38

Table 6. Reliability, Raw Score and Scaled Score Descriptive Statistics for MontCAS ELP 2008-2009 Test Forms by Grade (Continued)

Grades 3-5				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
C1	Listening	12	0.84	15	9.3	4.0	1.56	93.1	18.5	7.31
	Speaking	12	0.86	15	9.1	4.9	1.83	93.4	20.9	7.89
	Reading	12	0.70	15	8.3	3.3	1.80	88.3	15.9	8.70
	Writing	12	0.85	15	7.2	4.1	1.61	85.3	20.5	8.05
	Comprehen	12	0.86	27	15.6	6.2	2.28	91.0	17.5	6.47
	Total	12	0.94	60	33.9	14.1	3.56	378.8	30.3	7.66

Grade 3				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
C2	Listening	457	0.71	18	11.0	3.2	1.75	101.8	11.5	6.26
	Speaking	460	0.66	18	13.9	4.0	2.36	104.2	19.2	11.26
	Reading	459	0.75	18	10.0	3.7	1.88	101.3	12.7	6.39
	Writing	456	0.67	18	8.2	3.0	1.69	99.9	10.9	6.21
	Comprehen	460	0.81	36	20.9	6.0	2.66	101.1	10.1	4.44
	Total	460	0.84	72	42.9	10.5	4.19	401.6	18.0	7.22

Grade 4				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
C2	Listening	412	0.67	18	12.6	2.9	1.66	107.4	11.1	6.40
	Speaking	413	0.61	18	15.3	2.7	1.70	110.8	15.1	9.35
	Reading	413	0.79	18	11.6	3.9	1.79	107.3	14.3	6.52
	Writing	413	0.69	18	10.0	3.1	1.73	106.5	11.8	6.59
	Comprehen	413	0.83	36	24.2	6.0	2.50	106.7	10.6	4.40
	Total	413	0.86	72	49.6	9.4	3.48	413.6	17.9	6.60

Grade 5				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
C2	Listening	340	0.75	18	13.0	3.1	1.56	109.2	12.2	6.16
	Speaking	342	0.62	18	15.6	2.8	1.75	112.8	15.9	9.88
	Reading	341	0.80	18	12.7	3.8	1.70	111.3	14.5	6.55
	Writing	341	0.75	18	10.5	3.4	1.70	108.3	13.5	6.77
	Comprehen	341	0.85	36	25.7	6.0	2.35	109.5	11.3	4.45
	Total	342	0.87	72	51.6	10.2	3.63	418.0	19.8	7.07

Table 6. Reliability, Raw Score and Scaled Score Descriptive Statistics for MontCAS ELP 2008-2009 Test Forms by Grade (Continued)

Grades 6-8				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
D1	Listening	13	0.35	15	8.8	2.2	1.81	86.0	6.8	5.48
	Speaking	13	0.88	15	7.2	5.9	2.02	82.9	20.9	7.19
	Reading	13	0.45	15	8.3	2.4	1.78	82.5	6.6	4.90
	Writing	13	0.81	15	8.5	3.3	1.44	82.4	10.2	4.41
	Comprehen	13	0.59	29	16.2	3.9	2.52	84.1	5.9	3.78
	Total	13	0.84	60	32.8	11.4	4.54	364.9	17.4	6.90

Grade 6				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
D2	Listening	303	0.72	18	12.6	3.0	1.59	99.1	8.5	4.50
	Speaking	307	0.75	18	14.8	3.3	1.65	104.6	12.4	6.26
	Reading	307	0.70	20	9.7	3.7	2.04	97.9	9.1	4.96
	Writing	305	0.73	20	11.4	3.3	1.72	96.3	9.0	4.74
	Comprehen	307	0.80	38	22.1	6.1	2.72	98.1	7.5	3.36
	Total	307	0.86	76	48.3	10.3	3.81	396.2	14.2	5.23

Grade 7				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
D2	Listening	288	0.74	18	13.3	3.0	1.54	101.6	9.4	4.75
	Speaking	291	0.68	18	15.1	3.3	1.86	106.1	13.5	7.63
	Reading	288	0.72	20	10.7	3.8	2.01	100.4	9.1	4.83
	Writing	288	0.74	20	12.3	3.3	1.66	98.8	9.3	4.74
	Comprehen	289	0.82	38	24.0	6.0	2.58	100.5	7.8	3.35
	Total	291	0.85	76	51.0	10.7	4.09	399.8	17.8	6.82

Grade 8				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
D2	Listening	334	0.78	18	13.9	3.0	1.41	103.7	10.0	4.72
	Speaking	338	0.73	18	15.1	3.6	1.88	107.0	14.5	7.54
	Reading	336	0.74	20	12.0	3.8	1.97	103.3	9.8	5.02
	Writing	334	0.76	20	13.3	3.5	1.71	102.1	10.7	5.27
	Comprehen	336	0.84	38	25.8	6.3	2.47	102.9	8.6	3.40
	Total	338	0.88	76	54.0	11.6	3.98	404.9	20.1	6.94

Table 6. Reliability, Raw Score and Scaled Score Descriptive Statistics for MontCAS ELP 2008-2009 Test Forms by Grade (Continued)

Grades 9-12				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
E1	Listening	9	0.85	15	8.9	4.2	1.60	88.0	16.9	6.51
	Speaking	9	0.83	15	9.9	3.9	1.62	90.0	13.0	5.36
	Reading	9	0.71	15	9.0	3.2	1.75	84.6	13.7	7.39
	Writing	9	0.60	15	6.2	2.6	1.64	78.4	8.1	5.12
	Comprehen	9	0.87	28	16.2	6.6	2.35	86.7	15.5	5.51
	Total	9	0.92	60	34.0	12.2	3.48	371.6	15.9	4.53

Grade 9				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
E2	Listening	275	0.77	18	11.6	3.4	1.64	96.1	10.1	4.83
	Speaking	281	0.61	18	14.2	3.7	2.27	103.1	13.8	8.58
	Reading	280	0.73	20	12.3	3.6	1.86	98.0	8.6	4.51
	Writing	278	0.71	20	10.7	3.3	1.77	96.1	8.2	4.43
	Comprehen	280	0.82	38	23.7	6.2	2.60	96.5	7.7	3.24
	Total	281	0.87	76	48.5	10.6	3.88	395.3	11.6	4.25

Grade 10				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
E2	Listening	249	0.81	18	12.2	3.6	1.58	97.7	10.6	4.68
	Speaking	254	0.65	18	13.9	3.8	2.28	101.7	13.5	8.00
	Reading	252	0.75	20	12.8	3.6	1.79	99.3	9.1	4.54
	Writing	250	0.72	20	11.3	3.3	1.74	97.7	8.4	4.46
	Comprehen	252	0.86	38	24.8	6.6	2.45	97.9	8.6	3.20
	Total	254	0.86	76	49.6	11.7	4.37	396.4	13.0	4.83

Grade 11				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
E2	Listening	193	0.85	18	13.2	3.7	1.41	101.6	12.2	4.69
	Speaking	198	0.63	18	13.9	4.3	2.63	102.0	16.2	9.80
	Reading	195	0.81	20	13.7	3.8	1.66	101.6	9.5	4.16
	Writing	194	0.78	20	12.0	3.5	1.67	99.4	9.3	4.42
	Comprehen	195	0.90	38	26.8	6.9	2.23	101.0	9.6	3.07
	Total	198	0.91	76	52.0	13.3	4.02	399.5	16.6	5.00

Grade 12				Raw Scores				Scaled Scores		
Form	Language Domain	N	Alpha	Max	Mean	Std. Dev.	SEM	Mean	Std. Dev.	SEM
E2	Listening	189	0.80	18	12.9	3.5	1.56	100.4	11.5	5.14
	Speaking	191	0.62	18	14.6	3.6	2.22	104.7	14.0	8.64
	Reading	190	0.74	20	13.7	3.5	1.77	101.8	9.5	4.84
	Writing	190	0.71	20	12.0	3.3	1.75	99.5	8.7	4.67
	Comprehen	190	0.86	38	26.5	6.4	2.40	100.7	9.7	3.63
	Total	191	0.87	76	53.0	10.9	3.97	400.3	14.5	5.30

9. Validity of the MontCAS ELP

9.1 Content-related Validity. Validity of the MontCAS ELP begins with test content. The Introduction to the Mountain West Assessment Consortium Foundation Document, included as an appendix to the *2006-2007 MontCAS ELP Technical Report*, provides background information on the design of the assessment. Additional information on the development of the Mountain West Items is provided in Matthews (2007). All of the items on the 2008-2009 MontCAS ELP were developed as part of that Mountain West Consortium effort.

9.2 Construct and Criterion-related Validity. In addition to test design considerations, test results also bear on the content validity of the assessment. In very general terms, the distribution and range of scores within each grade span and grade level (Table 6) provide evidence that the MontCAS ELP can capture a range of abilities. And, Table 7 (page 29) provides information on the validity of the assessment showing intercorrelations among components of the test. Table 7 shows, by grade span for Level 2 forms, Pearson product moment correlations among scaled scores on each subtest (Listening, Speaking, Reading, Writing, and Comprehension). Correlations are not reported for subtests that share common items (e.g, Reading and Comprehension) nor are they reported for subtests and Total MontCAS ELP. The number below the correlation coefficient in each cell represents the number of students on which the correlation is based.

Table 7. Correlations Among Scaled Scores on Individual Language Domain Tests

Grade	K	1-2	3-5	6-8	9-12	
r	A	B2	C2	D2	E2	Avg.
L x S	0.64 402	0.33 846	0.32 1,209	0.24 925	0.19 906	0.34
L x R	0.53 402	0.59 845	0.51 1,208	0.55 924	0.60 906	0.56
L x W	0.29 402	0.54 843	0.48 1,207	0.50 923	0.52 903	0.47
S x R	0.47 402	0.28 849	0.27 1,213	0.31 931	0.27 917	0.32
S x W	0.27 402	0.35 846	0.32 1,210	0.28 927	0.21 912	0.29
S x C	0.63 402	0.36 851	0.34 1,214	0.31 932	0.30 917	0.39
R x W	0.49 402	0.69 847	0.62 1,209	0.63 926	0.60 912	0.61
W x C	0.30 402	0.69 848	0.65 1,210	0.65 927	0.63 912	0. 58
Avg.	0.45	0.48	0.44	0.43	0.42	

All of the correlation coefficients in Table 7 are significantly different from zero, indicating that the different subtests are measuring related abilities. Insofar as the language domain tests are measuring aspects of the same construct, English proficiency, performance in the different domains should be related. In addition, however, most of the coefficients are not high enough to suggest that the abilities measured by the individual domain tests are identical, reinforcing the assumption that language domain abilities are different aspects of overall English proficiency. The pattern of values in Table 7 is similar to that obtained in 2006-2007 and 2007-2008, suggesting that the internal structure of the tests across years is similar.

One valuable source of evidence relevant to the validity of a test is the relationship between scores on the test to those on another test measuring a similar or related construct. Student scores on the MontCAS Criterion-referenced Test (CRT) were provided for a group of students who had also been administered the MontCAS ELP. Data bearing on the relationship between those measures of student ability are presented next. Table 8 (page 30) shows student ability as measured by the English Language Arts (ELA) and the Math portion of the CRT compared to ability measured by the MontCAS ELP. On each test, ELA, Math, and ELP test, students are classified, based on their performance, in one of four categories: Novice, Nearing Proficient, Proficient, and Advanced. Each cell shows the number of students who were

classified in each category on each test. So, for example, of the 1,518 students who were classified as proficient on the MontCAS ELP test, 337 were classified as Novice on the ELA portion of the CRT, 665 as Nearing Proficient, 473 as Proficient and 43 as Advanced. Overall, the table shows that there is a positive relationship between performance on the CRT and ELP test. The distribution of scores on the ELA CRT is different from that on the Math CRT. More than half of the students administered the Math CRT scored at the Novice level, whereas the distribution was more evenly divided on the ELA CRT.

Table 8. Numbers of Students Scoring at Each Performance Level on the MontCAS ELP test and on MontCAS Criterion-referenced Tests in ELA and Math

		MontCAS ELP				
CRT		**Novice**	**Nearing Proficient**	**Proficient**	**Advanced**	**Total**
ELA	**Novice**	48	283	337	3	671
	Nearing Prof.	11	151	665	36	863
	Proficient	11	19	473	123	626
	Advanced	3	9	43	68	123
	Total	73	462	1518	230	2283
Math	**Novice**	47	364	792	38	1241
	Nearing Prof.	15	76	475	54	620
	Proficient	9	12	217	101	339
	Advanced	2	10	34	37	83
	Total	73	462	1518	230	2283

A quantitative measure of the relationship between performance on these two tests was generated in the following way. Each of the categories for each test was assigned a point value as follows: Novice = 1; Nearing Proficient = 2; Proficient = 3; and Advanced = 4. Using these point values, a set of paired scores was generated, one pair for each of the 2,283 students administered the ELP and ELA tests and another pair for each of the 2,283 students administered the ELP and Math tests. Because these numbers represent ordinal categories rather than a continuous variable, a Spearman R correlation was calculated. Each cell in Table 9 (page 31) shows a Spearman correlation coefficient and the numbers of pairs of scores on which the coefficient was based. Correlational data are only available for grades in which the CRT is administered (i.e., 3-8 and 10). Coefficients are shown for all grades, two grade clusters and grade 10. All correlation coefficients are significantly different from zero,

establishing that there is a positive relation between performance on the MontCAS ELP test and the CRT in both ELA and Math. In addition, the correlation coefficients are uniformly higher for the ELA test than for the Math test. Finally, the coefficients for ELA remain relatively stable over grade clusters whereas the Math correlation drops off significantly in Grade 10.

Table 9. Correlations between Performance Levels on the MontCAS ELP and ELA and Math CRT

MontCAS CRT	All Grades	Grade 3-5	Grades 6-8	Grade 10
ELA	0.49 2,283	0.44 1,185	0.51 875	0.41 223
Math	0.35 2,283	0.38 1,184	0.36 876	0.17 223

10. MontCAS ELP Performance by Year

Table 10 (page 32) shows results for both 2007 and 2008 by form and grade, thus allowing a comparison of performance in those two years. This table shows, for each language domain, comprehension, and total MontCAS ELP, the maximum raw score (**RS_{Max}**), number of students (**N**) administered the assessment, the average raw score (**RS_{Mean}**) and average scaled score (**SS_{Mean}**). Whereas changes in average scaled scores can be used to compare performance across years within a grade, raw scores cannot be compared because of the reduction in 2008 of the number of RS points per form.

There are several generalizations that can be made. First, there was a decline from 2007 to 2008 in the number of students tested in each grade. The one exception to this generalization is an increase in the number of students administered B1 forms in grades 1 and 2. Second, as in 2006 and 2007, a relatively small percentage of the students were administered the Level 1 forms and that percentage was lowest in the upper grades. Third, performance on the total test as indicated by Total MontCAS ELP scaled score was comparable (i.e., generally within a standard error of measurement) across years within a grade. Fourth, as noted in the *2007-2008 Technical Report*, average raw scores in Listening and Speaking on Level 2 forms in several grades (2, 5, 6-8, and 9-12) were close to the maximum correct suggesting that these two subtests were not appropriately challenging for the population being tested. With the change in Listening and Speaking test configuration in the current year, the average raw score was generally not as close to the maximum score as in the 2006-2007 administration.

Table 10. Performance on 2007 and 2008 MontCAS ELP Test Forms by Grade

Form	Language Domain	RS_{Max}	N	RS_{Mean}	SS_{Mean}	RS_{Max}	N	RS_{Mean}	SS_{Mean}
		2007				2008			
Kindergarten									
A	Listening	22	443	12.0	104.5	15	402	9.4	106.5
	Speaking	22	437	14.5	108.5	15	402	11.0	106.1
	Reading	36	439	11.7	80.2	27	402	11.7	83.7
	Writing	22	404	6.0	66.3	22	402	5.9	67.4
	Comprehen	29	443	12.5	101.7	18	402	9.8	104.5
	Total	102	450	42.8	379.0	79	402	38.0	383.0
Grade 1-2									
B1	Listening	22	55	14.1	95.2	15	109	10.2	97.5
	Speaking	22	55	12.6	93.1	15	109	9.1	101.9
	Reading	15	55	9.9	94.2	15	110	8.9	86.7
	Writing	15	55	8.7	96.9	15	109	6.9	90.7
	Comprehen	31	55	20.1	94.5	23	110	14.0	91.7
	Total	74	55	45.3	384.5	60	110	34.9	383.3
Grade 1									
B2	Listening	22	484	15.2	100.0	18	383	10.5	103.1
	Speaking	22	478	16.3	109.5	18	384	13.0	108.6
	Reading	20	484	11.0	94.3	18	383	9.4	97.1
	Writing	20	483	6.3	85.9	18	382	4.6	88.6
	Comprehen	39	489	23.9	96.0	36	385	19.3	99.7
	Total	84	490	48.0	391.0	72	386	37.2	398.5
Grade 2									
B2	Listening	22	479	18.7	118.4	18	464	13.2	114.8
	Speaking	22	469	18.4	118.9	18	468	14.4	116.3
	Reading	20	475	15.6	114.4	18	468	13.5	115.4
	Writing	20	478	12.2	117.6	18	466	9.2	111.6
	Comprehen	39	480	31.6	115.2	36	468	26.0	114.2
	Total	84	481	64.2	440.5	72	468	50.1	435.7

Table 10. Performance on 2007 and 2008 MontCAS ELP Test Forms by Grade (continued)

Form	Language Domain	RS_{Max}	N	RS Mean	SS Mean	RS_{Max}	N	RS Mean	SS Mean
		2007				2008			
Grade 3-5									
C1	Listening	22	20	14.2	96.9	15	12	9.3	93.1
	Speaking	22	20	14.9	94.1	15	12	9.1	93.4
	Reading	15	20	9.4	94.9	15	12	8.3	88.3
	Writing	15	20	9.0	96.4	15	12	7.2	85.3
	Comprehen	31	20	19.3	95.8	27	12	15.6	91.0
	Total	74	21	45.1	381.1	60	12	33.9	378.8
Grade 3									
C2	Listening	22	476	15.8	101.8	18	457	11.0	101.8
	Speaking	22	476	18.8	108.7	18	460	13.9	104.2
	Reading	20	476	10.8	99.8	18	459	10.0	101.3
	Writing	19	473	9.8	100.3	18	456	8.2	99.9
	Comprehen	39	481	23.9	100.1	36	460	20.9	101.1
	Total	83	482	54.4	401.6	72	460	42.9	401.6
Grade 4									
C2	Listening	22	440	17.3	107.9	18	412	12.6	107.4
	Speaking	22	439	19.5	112.7	18	413	15.3	110.8
	Reading	20	441	12.5	105.1	18	413	11.6	107.3
	Writing	19	441	11.7	108.6	18	413	10.0	106.5
	Comprehen	39	442	27.2	105.7	36	413	24.2	106.7
	Total	83	443	60.5	412.9	72	413	49.6	413.6
Grade 5									
C2	Listening	22	382	18.5	112.6	18	340	13.0	109.2
	Speaking	22	372	19.8	114.0	18	342	15.6	112.8
	Reading	20	382	14.4	111.3	18	341	12.7	111.3
	Writing	19	382	12.7	114.2	18	341	10.5	108.3
	Comprehen	39	383	30.1	111.2	36	341	25.7	109.5
	Total	83	383	64.7	421.9	72	342	51.6	418.0

Table 10. Performance on 2007 and 2008 MontCAS ELP Test Forms by Grade (continued)

Form	Language Domain	RS_{Max}	N	RS_{Mean}	SS_{Mean}	RS_{Max}	N	RS_{Mean}	SS_{Mean}
		2007				2008			
Grade 6-8									
D1	Listening	22	30	16.6	98.7	15	13	8.8	86.0
	Speaking	22	29	15.9	95.7	15	13	7.2	82.9
	Reading	15	30	10.4	93.9	15	13	8.3	82.5
	Writing	15	30	11.3	100.3	15	13	8.5	82.4
	Comprehen	32	30	23.1	96.0	29	13	16.2	84.1
	Total	74	30	53.6	389.9	60	13	32.8	364.9
Grade 6									
D2	Listening	22	405	17.4	100.3	18	303	12.6	99.1
	Speaking	22	405	19.0	104.8	18	307	14.8	104.6
	Reading	24	408	13.3	99.1	20	307	9.7	97.9
	Writing	20	406	12.2	100.9	20	305	11.4	96.3
	Comprehen	43	408	28.8	98.4	38	307	22.1	98.1
	Total	88	410	61.2	399.6	76	307	48.3	396.2
Grade 7									
D2	Listening	22	418	18.0	102.6	18	288	13.3	101.6
	Speaking	22	419	18.9	106.2	18	291	15.1	106.1
	Reading	24	422	14.1	101.0	20	288	10.7	100.4
	Writing	20	419	12.7	102.4	20	288	12.3	98.8
	Comprehen	44	422	30.2	100.2	38	289	24.0	100.5
	Total	88	422	63.4	402.9	76	291	51.0	399.8
Grade 8									
D2	Listening	22	400	18.9	105.7	18	334	13.9	103.7
	Speaking	22	401	19.6	108.8	18	338	15.1	107.0
	Reading	24	405	15.4	103.5	20	336	12.0	103.3
	Writing	20	405	13.9	106.6	20	334	13.3	102.1
	Comprehen	44	405	32.2	102.9	38	336	25.8	102.9
	Total	88	407	66.9	408.4	76	338	54.0	404.9

Table 10. Performance on 2007 and 2008 MontCAS ELP Test Forms by Grade (continued)

Form	Language Domain	RS_{Max}	N	RS Mean	SS Mean	RS_{Max}	N	RS Mean	SS Mean
		2007				2008			
Grade 9-12									
E1	Listening	22	3	9.0	82.3	15	9	8.9	88.0
	Speaking	22	3	6.0	75.3	15	9	9.9	90.0
	Reading	15	3	8.0	86.3	15	9	9.0	84.6
	Writing	15	3	5.7	84.3	15	9	6.2	78.4
	Comprehen	32	3	13.3	81.7	28	9	16.2	86.7
	Total	74	3	28.7	366.3	60	9	34.0	371.6
Grade 9									
E2	Listening	22	494	16.0	98.3	18	275	11.6	96.1
	Speaking	22	490	17.6	102.4	18	281	14.2	103.1
	Reading	25	496	14.9	97.6	20	280	12.3	98.0
	Writing	20	496	11.1	98.6	20	278	10.7	96.1
	Comprehen	44	498	29.1	97.0	38	280	23.7	96.5
	Total	89	502	58.6	395.8	76	281	48.5	395.3
Grade 10									
E2	Listening	22	309	17.3	101.9	18	249	12.2	97.7
	Speaking	22	307	18.6	105.9	18	254	13.9	101.7
	Reading	25	310	16.5	101.2	20	252	12.8	99.3
	Writing	20	307	12.6	103.2	20	250	11.3	97.7
	Comprehen	45	312	31.9	100.5	38	252	24.8	97.9
	Total	89	314	63.9	402.0	76	254	49.6	396.4
Grade 11									
E2	Listening	22	311	17.7	103.4	18	193	13.2	101.6
	Speaking	22	320	18.8	106.5	18	198	13.9	102.0
	Reading	25	307	17.2	102.7	20	195	13.7	101.6
	Writing	20	306	13.0	104.6	20	194	12.0	99.4
	Comprehen	45	312	32.8	101.8	38	195	26.8	101.0
	Total	89	323	64.3	402.8	76	198	52.0	399.5
Grade 12									
E2	Listening	22	256	18.5	106.2	18	189	12.9	100.4
	Speaking	22	257	19.1	107.8	18	191	14.6	104.7
	Reading	25	258	17.2	102.8	20	190	13.7	101.8
	Writing	20	258	13.4	106.3	20	190	12.0	99.5
	Comprehen	45	258	33.8	103.2	38	190	26.5	100.7
	Total	89	262	66.9	406.3	76	191	53.0	400.3

Performance on MontCAS ELP 2007 and MontCAS ELP 2008 is summarized in Table 11, below. This table shows the percent of students in each Total MontCAS ELP Proficiency category by grade (N=Novice, NP=Nearing Proficiency, P=Proficient, A=Advanced). This table is not from a matched sample and includes all students tested in Fall 2007 and all tested in Fall 2008.

Table 11. Total MontCAS ELP Level by Grade in 2007 and 2008

Grade	Percent in each Proficiency Category								
	2007					2008			
	N	NP	P	A		N	NP	P	A
K	26	42	26	6		24	41	26	9
1	8	20	56	16		5	16	61	18
2	4	14	62	20		2	12	73	13
3	2	14	61	23		2	11	67	19
4	4	19	57	20		2	16	68	15
5	7	21	65	7		8	18	71	3
6	3	20	59	19		1	30	57	11
7	3	20	64	14		3	24	65	9
8	3	14	79	5		3	18	76	3
9	3	37	56	4		2	40	56	1
10	3	27	66	5		5	37	57	1
11	7	28	63	2		7	38	55	0
12	5	25	66	4		2	46	53	0
K-12	6	23	59	12		5	24	61	10

As in previous administrations of the MontCAS ELP, the percent in each proficiency category in each grade remains fairly stable from one year to the next. The most consistent change from 2007 to 2008 is a decrease in the percent in the Advanced category, shown in 11 of the 13 grades. One interpretation of this change is that many of the most proficient students tested in 2007 were not tested in 2008. This interpretation is consistent with the overall changes in numbers tested. There were approximately 1,000 fewer students tested in 2008 than in 2007.

Table 12 (page 38) shows a summary of MontCAS ELP Growth Reports by grade. Whereas Tables 10 and 11, comparing performance from 2007 to 2008, do not represent a matched sample (i.e., students who were tested in both years) Table 12 table represents the performance of students who were tested in both 2007 and 2008 and whose results were matched. Of 4,073 students in grades 1 through 12 tested in Fall 2008, test results for 2,715 or 66.7% were matched to the previous year (students in Kindergarten in 2008 were excluded since they were not tested in the previous year). Each cell in the table shows the number and percent of students by grade. Table 12 summarizes three categories of change in proficiency levels from 2007 to 2008. The "declining" category shows the number and percent of students whose proficiency level declined by one or more levels from 2007 to 2008. The "maintaining" category represents the number and percent of students who stayed at the same proficiency level, and the "gaining" category shows the number and percent that either remained at the advanced level or gained in proficiency by one or more levels.

In every grade except grade 1, the largest percentage of students fell into the "maintaining" category. The highest percentages in the "gaining" category were in the lower grades and the highest percentages in the "declining" category were in the upper grades.

For comparison purposes, Table 13 (page 39) shows a summary of MontCAS ELP Growth Reports for those students who were tested in Fall 2006 and Fall 2007. Of the 5,028 students in grades 1 through 12 tested in Fall 2007, test results for 3,291 or 65.5% were matched to the previous year. From Fall 2006 to Fall 2007 and from Fall 2007 to Fall 2008, the percentages matched were very similar.

In addition, comparing the final row in Tables 12 and 13, the percentages over all grades in each of the three categories were very similar.

Table 12. Summary of 2007 to 2008 Growth Reports

Grade	Declining	Maintaining	Gaining
1	**8** (2.7%)	**71** (24.4%)	**212** (72.9%)
2	**26** (8.5%)	**165** (53.7)	**116** (37.8%)
3	**46** (15.3%)	**167** (55.5%)	**88** (29.2%)
4	**33** (11.3%)	**198** (68.0%)	**60** (20.6%)
5	**50** (19.5%)	**161** (62.9%)	**45** (17.6%)
6	**28** (12.8%)	**128** (58.7%)	**62** (28.4%)
7	**25** (11.8%)	**143** (67.8%)	**43** (20.4%)
8	**37** (14.6%)	**180** (71.1%)	**36** (14.2%)
9	**46** (28.8%)	**99** (61.9%)	**15** (9.4%)
10	**27** (15.1%)	**118** (65.9%)	**34** (19.0%)
11	**27** (23.7%)	**78** (68.4%)	**9** (7.9%)
12	**17** (12.7%)	**100** (74.6%)	**17** (12.7%)
1-12	**370** (13.6%)	**1608** (59.2%)	**737** (27.1%)

Table 13. Summary of 2006 to 2007 Growth Reports

Grade	Declining	Maintaining	Gaining
1	**29** (9.1%)	**106** (33.4%)	**182** (57.4%)
2	**27** (7.9%)	**173** (50.7)	**141** (41.3%)
3	**28** (9.4%)	**168** (56.4%)	**102** (34.2%)
4	**43** (13.9%)	**166** (53.5%)	**101** (32.6%)
5	**54** (19.4%)	**178** (63.8%)	**47** (16.8%)
6	**30** (10.1%)	**162** (54.7%)	**104** (35.1%)
7	**28** (9.2%)	**187** (61.7%)	**88** (29.0%)
8	**30** (10.1%)	**198** (66.4%)	**70** (23.5%)
9	**40** (15.7%)	**172** (67.7%)	**42** (16.5%)
10	**19** (9.4%)	**149** (73.8%)	**34** (16.8%)
11	**31** (15.5%)	**147** (73.5%)	**22** (11.0%)
12	**19** (9.8%)	**130** (67.4%)	**44** (22.8%)
1-12	**378** (11.5%)	**1936** (58.8%)	**977** (29.7%)

References

Linacre, J. M. & Wright, B. D. (2005). *A user's guide to WINSTEPS: Rasch-model computer program* (v. 3.57). Chicago, IL: MESA Press.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

Matthews, G. (2007). Developing the Mountain West assessment. In J. Abedi (ed.). *English language proficiency in the nation (pp.* 33-45). Davis, CA: University of California, School of Education.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Appendix A: Item Difficulty and Discrimination data

Table A.1. Grade K: Form A

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88066	1	Listening	CR	Core	1	402	34	56				11	0.56	0.48
88064	2	Listening	CR	Core	1	402	41	45				14	0.45	0.40
88069	3	Listening	MC	Core	1	402		88	7	3		1	0.88	0.25
88072	4	Listening	MC	Core	1	402		9	87	3		1	0.87	0.27
88070	5	Listening	MC	Core	1	402		8	11	80		1	0.80	0.22
88067	6	Listening	CR	Core	1	402	32	67				2	0.67	0.39
88071	7	Listening	CR	Core	1	402	16	83				1	0.83	0.34
88068	8	Listening	CR	Core	1	402	24	62				14	0.62	0.46
88065	9	Listening	CR	Core	1	402	29	59				12	0.59	0.45
8038003	10	Listening	CR	Core	1	402	26	62				12	0.62	0.45
8038004	11	Listening	CR	Core	1	402	47	39				15	0.39	0.33
8040001	12	Listening	CR	Core	1	402	30	59				11	0.59	0.37
8040002	13	Listening	CR	Core	1	402	45	36				19	0.36	0.35
8040003	14	Listening	CR	Core	1	402	15	75				9	0.75	0.37
8040005	15	Listening	CR	Core	1	402	39	46				15	0.46	0.37
88125	1	Speaking	CR	Core	1	402	5	92				3	0.92	0.26
88303	2	Speaking	CR	Core	1	402	6	91				2	0.91	0.26
88306	3	Speaking	CR	Core	1	402	11	84				5	0.84	0.36
88407	4	Speaking	CR	Core	1	402	34	54				12	0.54	0.39
88122	5	Speaking	CR	Core	1	402	9	89				1	0.89	0.25
88308	6	Speaking	CR	Core	1	402	7	91				2	0.91	0.32
88309	7	Speaking	CR	Core	1	402	15	82				4	0.82	0.34
88413	8	Speaking	CR	Core	2	402	7	33	58			2	0.75	0.47
88414	9	Speaking	CR	Core	4	402	4	18	28	30	11	8	0.52	0.48
88130	10	Speaking	CR	Core	4	402	8	26	27	20	5	14	0.40	0.53
88104	1	Reading	MC	Core	1	402		77	7	14		2	0.77	0.28
88281	2	Reading	MC	Core	1	402		26	64	8		2	0.64	0.33
88084	3	Reading	CR	Core	1	402	32	61				7	0.61	0.45
88288	4	Reading	CR	Core	1	402	30	61				9	0.61	0.48
88095	5	Reading	MC	Core	1	402		14	13	65		7	0.65	0.44
88291	6	Reading	MC	Core	1	402		13	67	9		10	0.67	0.34
88292	7	Reading	MC	Core	1	402		9	7	74		10	0.74	0.44
88098	8	Reading	CR	Core	1	402	28	56				16	0.56	0.51
88282	9	Reading	CR	Core	1	402	39	42				19	0.42	0.44
88283	10	Reading	CR	Core	1	402	32	51				17	0.51	0.47
88085	11	Reading	CR	Core	1	402	30	49				22	0.49	0.47
88284	12	Reading	CR	Core	1	402	35	42				24	0.42	0.47
88285	13	Reading	CR	Core	1	402	34	41				24	0.41	0.45
88286	14	Reading	CR	Core	1	402	30	41				29	0.41	0.64
88093	15	Reading	CR	Core	1	402	29	39				32	0.39	0.60
88287	16	Reading	CR	Core	1	402	23	43				34	0.43	0.66
88089	17	Reading	CR	Core	1	402	24	40				36	0.40	0.54
88289	18	Reading	CR	Core	1	402	42	19				39	0.19	0.42

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.1. Grade K: Form A (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88290	19	Reading	CR	Core	1	402	40	20				40	0.20	0.46
88100	20	Reading	CR	Core	1	402	33	26				41	0.26	0.54
88295	21	Reading	CR	Core	1	402	33	24				44	0.24	0.52
88296	22	Reading	CR	Core	1	402	31	24				45	0.24	0.54
88087	23	Reading	MC	Core	1	402		33	11	8		47	0.33	0.42
88103	24	Reading	MC	Core	1	402		32	8	10		49	0.32	0.42
88299	25	Reading	MC	Core	1	402		17	22	12		49	0.22	0.27
88119	26	Reading	MC	Core	1	402		26	12	11		50	0.26	0.32
88301	27	Reading	MC	Core	1	402		10	11	27		51	0.27	0.34
88448	1	Writing	SR1	Core	1	402	20	22	12	38		8	0.50	0.18
88449	2	Writing	SR1	Core	1	402	13	27	34	19		7	0.53	0.44
88450	3	Writing	SR1	Core	1	402	17	29	33	14		7	0.48	0.47
88451	4	Writing	SR1	Core	1	402	13	28	33	19		8	0.52	0.47
88452	5	Writing	SR1	Core	1	402	5	13	28	46		7	0.74	0.40
88453	6	Writing	SR1	Core	1	402	24	37	21	7		11	0.28	0.46
88454	7	Writing	SR1	Core	1	402	26	33	20	13		8	0.33	0.49
88455	8	Writing	SR1	Core	1	402	59	19	10	4		8	0.14	0.49
88456	9	Writing	SR1	Core	1	402	33	29	20	9		8	0.29	0.57
88457	10	Writing	SR1	Core	1	402	51	24	11	4		9	0.16	0.48
88458	11	Writing	SR1	Core	1	402	68	16	3	3		9	0.06	0.40
88459	12	Writing	SR1	Core	1	402	20	28	24	19		8	0.44	0.56
88460	13	Writing	SR1	Core	1	402	18	28	29	17		8	0.46	0.56
88461	14	Writing	SR1	Core	1	402	31	27	18	15		8	0.34	0.56
88462	15	Writing	SR1	Core	1	402	63	19	7	2		9	0.09	0.43
88463	16	Writing	SR1	Core	1	402	66	15	7	3		9	0.10	0.45
88464	17	Writing	SR1	Core	1	402	71	14	4	2		9	0.06	0.39
88465	18	Writing	SR1	Core	1	402	58	22	8	3		9	0.12	0.45
88466	19	Writing	SR1	Core	1	402	76	11	3	1		9	0.04	0.32
88467	20	Writing	SR1	Core	1	402	61	19	7	4		9	0.11	0.45
88468	21	Writing	SR1	Core	1	402	71	13	5	2		9	0.07	0.36
88469	22	Writing	SR1	Core	1	402	74	11	4	2		9	0.06	0.40

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.2. Grades 1-2: Form B1

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88002	1	Listening	MC	Core	1	110		7	2	87		3	0.87	0.18
88003	2	Listening	MC	Core	1	110		89	5	2		4	0.89	0.19
88004	3	Listening	MC	Core	1	110		87	5	5		3	0.87	0.13
88006	4	Listening	MC	Core	1	110		85	5	7		3	0.85	0.06
88417	5	Listening	MC	Core	1	110		79	15	2		4	0.79	0.18
88005	6	Listening	MC	Core	1	110		25	5	67		4	0.67	0.31
88009	7	Listening	MC	Core	1	110		15	74	9		3	0.74	0.48
8002001	8	Listening	MC	Core	1	110		34	46	15		5	0.46	0.40
8002002	9	Listening	MC	Core	1	110		50	22	24		5	0.50	0.20
8002003	10	Listening	MC	Core	1	110		71	21	5		4	0.71	0.27
8002004	11	Listening	MC	Core	1	110		8	5	84		4	0.84	0.22
8043001	12	Listening	MC	Core	1	110		74	8	13		5	0.74	0.34
8043002	13	Listening	MC	Core	1	110		43	30	23		5	0.30	0.27
8043004	14	Listening	MC	Core	1	110		22	39	34		5	0.34	0.30
8043005	15	Listening	MC	Core	1	110		20	51	25		5	0.51	0.47
88321	1	Speaking	CR	Core	1	110	15	81				5	0.81	0.30
88323	2	Speaking	CR	Core	1	110	18	75				6	0.75	0.35
88324	3	Speaking	CR	Core	1	110	15	77				7	0.77	0.33
88011	4	Speaking	CR	Core	1	110	30	65				5	0.65	0.57
88012	5	Speaking	CR	Core	1	110	39	54				7	0.54	0.42
88319	6	Speaking	CR	Core	1	110	14	79				7	0.79	0.46
88400	7	Speaking	CR	Core	1	110	28	63				9	0.63	0.48
88401	8	Speaking	CR	Core	2	110	21	36	37			5	0.55	0.53
88021	9	Speaking	CR	Core	2	110	14	45	37			5	0.60	0.44
88326	10	Speaking	CR	Core	4	110	10	28	25	21	10	5	0.45	0.58
88419	1	Reading	MC	Core	1	110		87	3	9		1	0.87	0.42
88026	2	Reading	MC	Core	1	110		6	1	92		1	0.92	0.37
88312	3	Reading	MC	Core	1	110		1	95	4		1	0.95	0.30
88029	4	Reading	MC	Core	1	110		11	51	37		1	0.51	0.01
88028	5	Reading	MC	Core	1	110		37	50	12		1	0.37	0.29
88313	6	Reading	MC	Core	1	110		77	8	14		1	0.77	0.44
88423	7	Reading	MC	Core	1	110		23	13	62		3	0.62	0.37
88424	8	Reading	MC	Core	1	110		73	15	11		1	0.73	0.28
88314	9	Reading	MC	Core	1	110		31	27	35		7	0.35	0.45
88034	10	Reading	MC	Core	1	110		63	16	15		5	0.63	0.31
88035	11	Reading	MC	Core	1	110		19	36	41		4	0.41	0.32
88316	12	Reading	MC	Core	1	110		26	45	25		3	0.45	0.22
8044001	13	Reading	MC	Core	1	110		32	40	25		2	0.32	0.13
8044002	14	Reading	MC	Core	1	110		23	22	47		8	0.47	0.14
8044003	15	Reading	MC	Core	1	110		23	51	22		4	0.51	0.22
88043	1	Writing	CR	Core	1	110	27	69				4	0.69	0.44
88044	2	Writing	CR	Core	1	110	10	85				5	0.85	0.39
88045	3	Writing	CR	Core	1	110	27	71				2	0.71	0.43
88046	4	Writing	CR	Core	1	110	35	62				3	0.62	0.51

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.2. Grades 1-2: Form B1 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88047	5	Writing	CR	Core	1	110	61	35				4	0.35	0.33
88330	6	Writing	CR	Core	1	110	53	43				5	0.43	0.51
88402	7	Writing	CR	Core	1	110	52	45				4	0.45	0.40
88331	8	Writing	CR	Core	1	110	50	45				5	0.45	0.49
88051	9	Writing	CR	Core	1	110	46	48				5	0.48	0.49
88052	10	Writing	CR	Core	1	110	57	34				9	0.34	0.36
88053	11	Writing	CR	Core	1	110	59	33				8	0.33	0.41
88054	12	Writing	CR	Core	2	110	41	38	13			8	0.32	0.59
88055	13	Writing	CR	Core	2	110	42	35	8			15	0.26	0.62

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.3. Grades 1-2: Form B2

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88417	1	Listening	MC	Core	1	854		91	7	1		2	0.91	0.24
88005	2	Listening	MC	Core	1	854		14	6	78		2	0.78	0.32
88009	3	Listening	MC	Core	1	854		9	83	7		1	0.83	0.27
88008	4	Listening	MC	Core	1	854		11	55	28		3	0.55	0.25
8002001	5	Listening	MC	Core	1	854		22	67	8		2	0.67	0.34
8002002	6	Listening	MC	Core	1	854		67	10	19		3	0.67	0.42
8002003	7	Listening	MC	Core	1	854		85	9	3		3	0.85	0.32
8002004	8	Listening	MC	Core	1	854		3	3	92		2	0.92	0.31
8043001	9	Listening	MC	Core	1	854		82	8	9		1	0.82	0.38
8043002	10	Listening	MC	Core	1	854		30	46	21		4	0.46	0.27
8043003	11	Listening	MC	Core	1	854		32	54	10		3	0.54	0.32
8043004	12	Listening	MC	Core	1	854		12	18	68		2	0.68	0.32
8043005	13	Listening	MC	Core	1	854		6	76	13		4	0.76	0.40
8003001	14	Listening	MC	Core	1	854		38	37	16		9	0.37	0.17
8003002	15	Listening	MC	Core	1	854		29	4	60		6	0.60	0.19
8003003	16	Listening	MC	Core	1	854		25	45	21		9	0.45	0.14
8003004	17	Listening	MC	Core	1	854		63	22	7		6	0.63	0.15
8003005	18	Listening	MC	Core	1	854		26	26	38		10	0.38	0.20
88016	1	Speaking	CR	Core	1	854	2	95				3	0.95	0.32
88015	2	Speaking	CR	Core	1	854	6	91				3	0.91	0.30
88018	3	Speaking	CR	Core	1	854	5	92				3	0.92	0.30
88019	4	Speaking	CR	Core	1	854	5	92				3	0.92	0.31
88014	5	Speaking	CR	Core	1	854	7	89				4	0.89	0.42
88013	6	Speaking	CR	Core	1	854	5	91				4	0.91	0.36
88401	7	Speaking	CR	Core	2	854	5	27	64			4	0.77	0.53
88021	8	Speaking	CR	Core	2	854	4	26	63			6	0.76	0.43
88022	9	Speaking	CR	Core	4	854	3	10	22	31	29	5	0.66	0.48
88023	10	Speaking	CR	Core	4	854	2	14	24	30	25	6	0.63	0.54
88036	1	Reading	MC	Core	1	854		9	84	4		3	0.84	0.29
88035	2	Reading	MC	Core	1	854		8	9	81		2	0.81	0.42
88421	3	Reading	MC	Core	1	854		2	87	9		1	0.87	0.20
88422	4	Reading	MC	Core	1	854		85	10	4		1	0.85	0.23
88316	5	Reading	MC	Core	1	854		6	78	15		2	0.78	0.32
88040	6	Reading	MC	Core	1	854		86	7	5		2	0.86	0.37
8005001	7	Reading	MC	Core	1	854		57	14	26		2	0.57	0.28
8005002	8	Reading	MC	Core	1	854		16	71	9		4	0.71	0.44
8005003	9	Reading	MC	Core	1	854		18	63	14		5	0.63	0.36
8005004	10	Reading	MC	Core	1	854		20	18	58		3	0.58	0.36
8006001	11	Reading	MC	Core	1	854		18	56	18		8	0.56	0.29
8006002	12	Reading	MC	Core	1	854		9	9	75		6	0.75	0.47
8006003	13	Reading	MC	Core	1	854		42	33	18		7	0.42	0.24
8006005	14	Reading	MC	Core	1	854		27	24	41		8	0.41	0.24
8046006	15	Reading	MC	Core	1	854		27	53	11		10	0.53	0.41
8046003	16	Reading	MC	Core	1	854		28	20	42		10	0.42	0.37
8046004	17	Reading	MC	Core	1	854		15	60	14		11	0.60	0.44

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.3. Grades 1-2: Form B2 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
8046005	18	Reading	MC	Core	1	854		18	43	28		12	0.43	0.33
88056	1	Writing	CR	Core	1	854	29	69				3	0.69	0.42
88335	2	Writing	CR	Core	1	854	43	54				4	0.54	0.35
88052	3	Writing	CR	Core	1	854	57	41				3	0.41	0.24
88336	4	Writing	CR	Core	1	854	46	52				2	0.52	0.33
88058	5	Writing	CR	Core	1	854	57	39				4	0.39	0.43
88049	6	Writing	CR	Core	1	854	33	62				5	0.62	0.45
88060	7	Writing	CR	Core	1	854	46	51				2	0.51	0.55
88057	8	Writing	CR	Core	1	854	58	39				4	0.39	0.53
88061	9	Writing	CR	Core	2	854	14	52	31			2	0.57	0.57
88062	10	Writing	CR	Core	4	854	21	41	27	3	0	7	0.27	0.63
88063	11	Writing	CR	Core	4	854	28	31	20	4	0	16	0.21	0.54

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.4. Grades 3-5: Form C1

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88146	1	Listening	MC	Core	1	12				8	92		0.92	0.43
88140	2	Listening	MC	Core	1	12		8	25	58	8		0.58	0.65
88408	3	Listening	MC	Core	1	12		42		17	42		0.42	0.43
88158	4	Listening	MC	Core	1	12		8	17	50	25		0.50	0.63
8010002	5	Listening	MC	Core	1	12		83	8		8		0.83	0.41
8010001	6	Listening	MC	Core	1	12		33	33	17	17		0.33	0.37
8010003	7	Listening	MC	Core	1	12		17	67	8	8		0.67	0.49
8010004	8	Listening	MC	Core	1	12			17	75	8		0.75	0.11
8047002	9	Listening	MC	Core	1	12		17	33	50			0.50	0.72
8047003	10	Listening	MC	Core	1	12		8	75	8	8		0.75	0.60
8047004	11	Listening	MC	Core	1	12		17	17		67		0.67	0.63
8012002	12	Listening	MC	Core	1	12				8	92		0.92	0.13
8012004	13	Listening	MC	Core	1	12		17	8	17	58		0.58	0.60
8012001	14	Listening	MC	Core	1	12		17	50	8	25		0.50	0.33
8012005	15	Listening	MC	Core	1	12		42	17	17	25		0.42	0.77
88428	1	Speaking	CR	Core	1	12		100					1.00	0.00
88144	2	Speaking	CR	Core	1	12	17	83					0.83	0.59
88157	3	Speaking	CR	Core	1	12	50	50					0.50	0.46
88343	4	Speaking	CR	Core	1	12	17	67				17	0.67	0.66
88344	5	Speaking	CR	Core	1	12	33	67					0.67	0.17
88345	6	Speaking	CR	Core	1	12		83				17	0.83	0.59
88154	7	Speaking	CR	Core	1	12	17	67				17	0.67	0.66
88347	8	Speaking	CR	Core	2	12	17	25	42			17	0.54	0.84
88143	9	Speaking	CR	Core	2	12	8	33	33			25	0.50	0.83
88148	10	Speaking	CR	Core	4	12	8	8	17	25	17	25	0.46	0.68
88168	1	Reading	MC	Core	1	12		75	8		17		0.75	0.01
88169	2	Reading	MC	Core	1	12		25		50	25		0.50	0.42
88174	3	Reading	MC	Core	1	12		83			17		0.83	0.14
88175	4	Reading	MC	Core	1	12			50		50		0.50	0.41
88187	5	Reading	MC	Core	1	12			17	58	25		0.58	0.27
88178	6	Reading	MC	Core	1	12		33	50		17		0.50	-0.17
8050001	7	Reading	MC	Core	1	12		17	67		17		0.67	0.27
8050002	8	Reading	MC	Core	1	12		25	8	58	8		0.58	0.66
8050003	9	Reading	MC	Core	1	12		58	17		25		0.58	0.65
8050004	10	Reading	MC	Core	1	12		17	17	8	58		0.58	0.06
8051001	11	Reading	MC	Core	1	12		17	17	25	42		0.25	0.24
8051002	12	Reading	MC	Core	1	12		8	17	58	17		0.58	0.42
8051003	13	Reading	MC	Core	1	12		58		33	8		0.58	0.42
8051004	14	Reading	MC	Core	1	12		17	8	25	50		0.50	0.13
8051005	15	Reading	MC	Core	1	12		33	33	33			0.33	0.38
88171	1	Writing	CR	Core	1	12	8	75				17	0.75	0.50
88164	2	Writing	CR	Core	1	12	25	58				17	0.58	0.77
88351	3	Writing	CR	Core	1	12	25	67				8	0.67	0.71
88349	4	Writing	CR	Core	1	12	25	58				17	0.58	0.24
88166	5	Writing	CR	Core	1	12		75				25	0.75	0.78

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.4. Grades 3-5: Form C1 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88167	6	Writing	MC	Core	1	12		17	50	8	17	8	0.50	0.56
88173	7	Writing	MC	Core	1	12		17	25		50	8	0.50	0.69
88353	8	Writing	MC	Core	1	12		25	50	17		8	0.50	0.28
88188	9	Writing	MC	Core	1	12		8		42	42	8	0.42	0.29
8015001	10	Writing	CR	Core	2	12	17	42	17			25	0.38	0.81
88170	11	Writing	CR	Core	4	12	8	25	33	8		25	0.29	0.61

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.5. Grades 3-5: Form C2

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88140	1	Listening	MC	Core	1	1,215		5	1	92	1	0	0.92	0.31
88408	2	Listening	MC	Core	1	1,215		3	12	11	73	1	0.73	0.29
88158	3	Listening	MC	Core	1	1,215		17	10	64	8	1	0.64	0.37
88399	4	Listening	MC	Core	1	1,215		90	4	3	1	1	0.90	0.32
89001	5	Listening	MC	Core	1	1,215		3	91	2	3	1	0.91	0.35
8049001	6	Listening	MC	Core	1	1,215		78	13	4	4	1	0.78	0.32
8049002	7	Listening	MC	Core	1	1,215		12	16	60	11	1	0.60	0.27
8049003	8	Listening	MC	Core	1	1,215		12	76	6	6	1	0.76	0.36
8049004	9	Listening	MC	Core	1	1,215		8	6	7	77	1	0.77	0.39
8049005	10	Listening	MC	Core	1	1,215		8	79	6	5	1	0.79	0.38
8011003	11	Listening	MC	Core	1	1,215		69	5	13	11	1	0.69	0.28
8011001	12	Listening	MC	Core	1	1,215		24	16	51	8	1	0.51	0.34
8011004	13	Listening	MC	Core	1	1,215		7	50	34	9	1	0.50	0.15
8011002	14	Listening	MC	Core	1	1,215		36	39	15	9	1	0.36	0.10
8048003	15	Listening	MC	Core	1	1,215		67	10	16	6	1	0.67	0.27
8048004	16	Listening	MC	Core	1	1,215		4	10	74	11	1	0.74	0.31
8048002	17	Listening	MC	Core	1	1,215		7	8	51	34	1	0.51	0.38
8048001	18	Listening	MC	Core	1	1,215		13	28	13	44	2	0.28	0.14
88142	1	Speaking	CR	Core	1	1,215	8	87				4	0.87	0.29
88163	2	Speaking	CR	Core	1	1,215	10	88				3	0.88	0.31
88344	3	Speaking	CR	Core	1	1,215	3	95				2	0.95	0.24
88156	4	Speaking	CR	Core	1	1,215	3	96				2	0.96	0.26
88155	5	Speaking	CR	Core	1	1,215	7	90				2	0.90	0.29
88154	6	Speaking	CR	Core	1	1,215	8	90				2	0.90	0.25
88141	7	Speaking	CR	Core	2	1,215	1	21	76			2	0.87	0.40
88143	8	Speaking	CR	Core	2	1,215	1	25	71			3	0.83	0.48
88148	9	Speaking	CR	Core	4	1,215	1	7	17	31	39	4	0.73	0.43
88150	10	Speaking	CR	Core	4	1,215	1	5	17	35	40	3	0.76	0.49
88186	1	Reading	MC	Core	1	1,215		68	7	6	17	0	0.68	0.29
88189	2	Reading	MC	Core	1	1,215		14	59	15	10	1	0.59	0.28
8050001	3	Reading	MC	Core	1	1,215		4	86	4	6	1	0.86	0.40
8050002	4	Reading	MC	Core	1	1,215		8	7	75	9	1	0.75	0.40
8050004	5	Reading	MC	Core	1	1,215		4	5	5	84	1	0.84	0.42
8018004	6	Reading	MC	Core	1	1,215		9	79	7	4	1	0.79	0.43
8018001	7	Reading	MC	Core	1	1,215		5	13	7	74	1	0.74	0.39
8018003	8	Reading	MC	Core	1	1,215		17	36	41	5	1	0.41	0.23
8018002	9	Reading	MC	Core	1	1,215		47	26	18	7	2	0.47	0.19
8017003	10	Reading	MC	Core	1	1,215		76	11	5	7	2	0.76	0.38
8017002	11	Reading	MC	Core	1	1,215		6	57	28	8	2	0.57	0.42
8017001	12	Reading	MC	Core	1	1,215		26	10	46	17	2	0.46	0.43
8017004	13	Reading	MC	Core	1	1,215		6	4	9	79	2	0.79	0.43
8053001	14	Reading	MC	Core	1	1,215		16	6	55	22	2	0.55	0.34
8053002	15	Reading	MC	Core	1	1,215		7	21	11	57	3	0.57	0.29
8053003	16	Reading	MC	Core	1	1,215		18	52	12	14	4	0.52	0.43
8053004	17	Reading	CR	Core	2	1,215	30	32	32			6	0.48	0.43

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.5. Grades 3-5: Form C2 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88352	1	Writing	MC	Core	1	1,215		73	13	9	4	1	0.73	0.30
88184	2	Writing	MC	Core	1	1,215		11	76	5	6	1	0.76	0.42
88181	3	Writing	MC	Core	1	1,215		12	6	4	76	1	0.76	0.45
88359	4	Writing	MC	Core	1	1,215		10	7	69	13	1	0.69	0.31
88353	5	Writing	MC	Core	1	1,215		6	74	12	7	1	0.74	0.34
88360	6	Writing	MC	Core	1	1,215		20	56	10	12	1	0.56	0.30
88398	7	Writing	MC	Core	1	1,215		9	16	66	7	1	0.66	0.33
88188	8	Writing	MC	Core	1	1,215		4	7	77	10	2	0.77	0.36
8019001	9	Writing	CR	Core	2	1,215	22	56	20			2	0.48	0.48
88179	10	Writing	CR	Core	4	1,215	6	17	56	16	2	3	0.46	0.52
88180	11	Writing	CR	Core	4	1,215	30	31	22	6	1	10	0.24	0.44

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.6. Grades 6-8: Form D1

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88200	1	Listening	MC	Core	1	13		69		23	8		0.69	0.13
88206	2	Listening	MC	Core	1	13		38	15	31	8	8	0.31	0.76
88207	3	Listening	MC	Core	1	13		23	46	23	8		0.46	-0.40
8022004	4	Listening	MC	Core	1	13		8	8	8	77		0.77	0.67
8022002	5	Listening	MC	Core	1	13		23	8	54	15		0.54	0.16
8022003	6	Listening	MC	Core	1	13		54	23	23			0.54	0.61
8022001	7	Listening	MC	Core	1	13		8	92				0.92	-0.24
8020002	8	Listening	MC	Core	1	13		8	46	23	23		0.46	-0.21
8020001	9	Listening	MC	Core	1	13		23	15	38	23		0.23	-0.53
8020003	10	Listening	MC	Core	1	13		8			92		0.92	-0.34
8020004	11	Listening	MC	Core	1	13		8	8	31	54		0.54	0.45
8056001	12	Listening	MC	Core	1	13		62	15	8	8	8	0.62	0.57
8056003	13	Listening	MC	Core	1	13		15	62	8	8	8	0.62	0.20
8056004	14	Listening	MC	Core	1	13		15	8	54	15	8	0.54	0.38
8056005	15	Listening	MC	Core	1	13		15	15	62		8	0.62	0.17
88195	1	Speaking	CR	Core	1	13	15	69				15	0.69	0.69
88211	2	Speaking	CR	Core	1	13	23	54				23	0.54	0.85
88362	3	Speaking	CR	Core	1	13	23	54				23	0.54	0.85
88191	4	Speaking	CR	Core	1	13	23	77					0.77	0.63
88196	5	Speaking	CR	Core	1	13	31	54				15	0.54	0.63
88365	6	Speaking	CR	Core	1	13	54	31				15	0.31	0.61
88212	7	Speaking	CR	Core	1	13	31	54				15	0.54	0.45
88210	8	Speaking	CR	Core	1	13	15	46				38	0.46	0.82
88209	9	Speaking	CR	Core	1	13	31	46				23	0.46	0.71
88213	10	Speaking	CR	Core	2	13	15	8	46			31	0.50	0.85
88192	11	Speaking	CR	Core	4	13	15	15	8	15	15	31	0.35	0.81
88217	1	Reading	MC	Core	1	13		8	69	8	15		0.69	0.66
88220	2	Reading	MC	Core	1	13				15	85		0.85	0.51
88226	3	Reading	MC	Core	1	13				92	8		0.92	-0.08
88227	4	Reading	MC	Core	1	13		46	15	15	23		0.46	-0.28
88218	5	Reading	MC	Core	1	13		23		23	54		0.54	0.25
8023005	6	Reading	MC	Core	1	13		54	15	8	23		0.54	0.26
8023001	7	Reading	MC	Core	1	13			77		23		0.77	0.54
8023002	8	Reading	MC	Core	1	13		23	15	54	8		0.54	0.35
8023003	9	Reading	MC	Core	1	13		23		31	46		0.31	0.17
8023004	10	Reading	MC	Core	1	13		8	77		15		0.77	-0.07
8058001	11	Reading	MC	Core	1	13		15	46		38		0.38	-0.06
8058002	12	Reading	MC	Core	1	13		23	8	62	8		0.62	0.11
8058003	13	Reading	MC	Core	1	13		38	15	15	31		0.38	0.56
8058004	14	Reading	MC	Core	1	13		23	15	38	23		0.23	-0.14
8058005	15	Reading	MC	Core	1	13		23	31	15	31		0.31	0.39
88372	1	Writing	CR	Core	1	13	15	77				8	0.77	0.37
88225	2	Writing	CR	Core	1	13	15	85					0.85	-0.31
88222	3	Writing	CR	Core	1	13	38	62					0.62	0.54
88373	4	Writing	MC	Core	1	13		85	8		8		0.85	0.19

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.6. Grades 6-8: Form D1 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88229	5	Writing	MC	Core	1	13		38	15	8	38		0.38	0.28
88228	6	Writing	MC	Core	1	13			15	23	62		0.62	0.41
88230	7	Writing	MC	Core	1	13		85	8	8			0.85	0.51
88438	8	Writing	MC	Core	1	13			31	69			0.69	0.52
88374	9	Writing	MC	Core	1	13			8	77	15		0.77	0.44
88215	10	Writing	CR	Core	2	13	23	62	8			8	0.38	0.79
88377	11	Writing	CR	Core	4	13	8	23	54			15	0.33	0.68

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.7. Grades 6-8: Form D2

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88206	1	Listening	MC	Core	1	936		10	3	78	7	1	0.78	0.38
88207	2	Listening	MC	Core	1	936		15	63	12	9	1	0.63	0.26
88202	3	Listening	MC	Core	1	936		85	4	6	3	2	0.85	0.31
88203	4	Listening	MC	Core	1	936		4	11	77	7	2	0.77	0.36
88405	5	Listening	MC	Core	1	936		43	53	2	0	1	0.53	0.25
8055001	6	Listening	MC	Core	1	936		11	76	3	7	1	0.76	0.27
8055002	7	Listening	MC	Core	1	936		86	5	2	6	1	0.86	0.33
8055003	8	Listening	MC	Core	1	936		2	5	82	10	1	0.82	0.32
8055004	9	Listening	MC	Core	1	936		1	1	2	94	1	0.94	0.32
8056001	10	Listening	MC	Core	1	936		88	4	3	4	1	0.88	0.37
8056002	11	Listening	MC	Core	1	936		17	29	5	47	1	0.47	0.18
8056003	12	Listening	MC	Core	1	936		8	84	3	3	1	0.84	0.34
8056005	13	Listening	MC	Core	1	936		9	6	72	11	1	0.72	0.41
8021005	14	Listening	MC	Core	1	936		12	16	14	56	2	0.56	0.29
8021002	15	Listening	MC	Core	1	936		69	10	17	2	2	0.69	0.36
8021001	16	Listening	MC	Core	1	936		4	8	73	13	2	0.73	0.29
8021003	17	Listening	MC	Core	1	936		12	11	61	15	2	0.61	0.19
8021004	18	Listening	MC	Core	1	936		6	71	11	11	2	0.71	0.43
88191	1	Speaking	CR	Core	1	936	3	96				1	0.96	0.26
88362	2	Speaking	CR	Core	1	936	8	91				1	0.91	0.29
88365	3	Speaking	CR	Core	1	936	4	95				1	0.95	0.25
88211	4	Speaking	CR	Core	1	936	3	95				1	0.95	0.32
88210	5	Speaking	CR	Core	1	936	10	89				1	0.89	0.26
88209	6	Speaking	CR	Core	1	936	16	81				4	0.81	0.31
88213	7	Speaking	CR	Core	2	936	2	18	79			1	0.88	0.43
88208	8	Speaking	CR	Core	2	936	2	21	75			2	0.86	0.40
88370	9	Speaking	CR	Core	4	936	2	7	17	30	40	4	0.73	0.46
88193	10	Speaking	CR	Core	4	936	1	4	14	29	49	2	0.79	0.46
88235	1	Reading	MC	Core	1	936		3	9	7	80	1	0.80	0.35
88233	2	Reading	MC	Core	1	936		2	5	82	10	1	0.82	0.33
8025004	3	Reading	MC	Core	1	936		67	16	10	5	1	0.67	0.33
8025002	4	Reading	MC	Core	1	936		29	17	11	42	1	0.42	0.35
8025001	5	Reading	MC	Core	1	936		24	9	15	51	1	0.51	0.33
8025003	6	Reading	MC	Core	1	936		10	64	10	14	1	0.64	0.29
8027005	7	Reading	MC	Core	1	936		17	11	67	4	1	0.67	0.44
8027003	8	Reading	MC	Core	1	936		13	4	75	7	1	0.75	0.29
8027001	9	Reading	MC	Core	1	936		39	18	6	35	1	0.35	0.16
8027002	10	Reading	MC	Core	1	936		23	7	20	48	1	0.48	0.35
8027004	11	Reading	CR	Core	2	936	32	64	3			1	0.35	0.42
8060001	12	Reading	MC	Core	1	936		6	34	55	3	2	0.55	0.26
8060002	13	Reading	MC	Core	1	936		33	30	16	19	2	0.33	0.30
8060003	14	Reading	MC	Core	1	936		23	60	8	6	2	0.60	0.43
8060004	15	Reading	MC	Core	1	936		18	10	11	59	2	0.59	0.30
8060005	16	Reading	CR	Core	4	936	14	18	36	13	16	3	0.48	0.52

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.7. Grades 6-8: Form D2 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88373	1	Writing	MC	Core	1	936		85	11	2	1	1	0.85	0.36
88376	2	Writing	MC	Core	1	936		3	10	9	78	1	0.78	0.31
88229	3	Writing	MC	Core	1	936		23	6	3	67	1	0.67	0.31
88214	4	Writing	MC	Core	1	936		1	94	4	1	1	0.94	0.36
88374	5	Writing	MC	Core	1	936		1	3	94	1	1	0.94	0.35
8028001	6	Writing	MC	Core	1	936		9	71	10	8	1	0.71	0.40
8028002	7	Writing	MC	Core	1	936		4	2	4	88	1	0.88	0.40
8028003	8	Writing	MC	Core	1	936		67	11	16	4	1	0.67	0.32
8028004	9	Writing	MC	Core	1	936		12	6	43	38	1	0.43	0.19
8028005	10	Writing	MC	Core	1	936		7	9	62	21	1	0.62	0.35
88231	11	Writing	CR	Core	2	936	19	57	23			1	0.51	0.50
88378	12	Writing	CR	Core	4	936	3	21	46	19	8	2	0.51	0.58
88379	13	Writing	CR	Core	4	936	6	29	39	15	4	7	0.42	0.53

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.8. Grades 9-12: Form E1

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88439	1	Listening	MC	Core	1	9			89		11		0.89	-0.24
88250	2	Listening	MC	Core	1	9		22	44	22	11		0.44	0.70
88251	3	Listening	MC	Core	1	9		11	44	44			0.44	0.63
88247	4	Listening	MC	Core	1	9			11		89		0.89	0.41
8031001	5	Listening	MC	Core	1	9			11	67	22		0.67	0.33
8031003	6	Listening	MC	Core	1	9		67	11		22		0.67	0.44
8031002	7	Listening	MC	Core	1	9		44	44	11			0.44	0.45
8031004	8	Listening	MC	Core	1	9				33	67		0.67	0.59
8030003	9	Listening	MC	Core	1	9		33	22	33	11		0.33	0.63
8030001	10	Listening	MC	Core	1	9		11	11	11	67		0.67	0.40
8030004	11	Listening	MC	Core	1	9		22	11	33	33		0.33	0.78
8030002	12	Listening	MC	Core	1	9			78		22		0.78	0.63
8063002	13	Listening	MC	Core	1	9		33	67				0.67	0.37
8063003	14	Listening	MC	Core	1	9		44		56			0.44	0.49
8063005	15	Listening	MC	Core	1	9		11	11	56	22		0.56	0.45
88257	1	Speaking	CR	Core	1	9	11	89					0.89	0.41
88381	2	Speaking	CR	Core	1	9		100					1.00	0.00
88382	3	Speaking	CR	Core	1	9	11	89					0.89	0.13
88236	4	Speaking	CR	Core	1	9	44	44				11	0.44	0.74
88239	5	Speaking	CR	Core	1	9	33	44				22	0.44	0.74
88252	6	Speaking	CR	Core	1	9	22	78					0.78	0.41
88253	7	Speaking	CR	Core	1	9	33	67					0.67	0.65
88388	8	Speaking	CR	Core	2	9			67			33	0.67	0.77
88258	9	Speaking	CR	Core	2	9		33	67				0.83	0.65
88238	10	Speaking	CR	Core	4	9	22		67		11		0.44	0.66
88260	1	Reading	MC	Core	1	9					100		1.00	0.00
88261	2	Reading	MC	Core	1	9		67			33		0.67	0.52
88271	3	Reading	MC	Core	1	9			67	11	22		0.67	0.40
88443	4	Reading	MC	Core	1	9			33	44	22		0.33	-0.26
8032003	5	Reading	MC	Core	1	9		67	11	11	11		0.67	0.40
8032002	6	Reading	MC	Core	1	9		11	78		11		0.78	0.34
8032001	7	Reading	MC	Core	1	9		56		22	22		0.56	0.35
8032004	8	Reading	MC	Core	1	9			56		44		0.44	-0.06
8064001	9	Reading	MC	Core	1	9				56	44		0.56	0.15
8064005	10	Reading	MC	Core	1	9		11	78	11			0.78	0.63
8064003	11	Reading	MC	Core	1	9		56	33		11		0.56	0.66
8066001	12	Reading	MC	Core	1	9		44	22	33			0.33	0.83
8066002	13	Reading	MC	Core	1	9		44	33	11	11		0.44	0.49
8066003	14	Reading	MC	Core	1	9		11		67	22		0.67	0.44
8066004	15	Reading	MC	Core	1	9		22	56	11	11		0.56	0.45
88262	1	Writing	CR	Core	1	9	44	56					0.56	0.66
88390	2	Writing	CR	Core	1	9	44	56					0.56	0.59
88275	3	Writing	MC	Core	1	9			11	89			0.89	0.31
88268	4	Writing	MC	Core	1	9		78		22			0.78	0.13
88444	5	Writing	MC	Core	1	9		33	11	11	44		0.44	-0.18

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.8. Grades 9-12: Form E1 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88266	6	Writing	MC	Core	1	9		33	11	22	33		0.22	0.24
88267	7	Writing	MC	Core	1	9		56			44		0.44	0.31
88264	8	Writing	MC	Core	1	9		11	22	22	44		0.22	0.32
88392	9	Writing	MC	Core	1	9		22	44	33			0.44	0.02
88274	10	Writing	CR	Core	2	9	56	44					0.22	0.05
88265	11	Writing	CR	Core	4	9	22	33	44				0.31	0.85

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.9. Grades 9-12: Form E2

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
88251	1	Listening	MC	Core	1	924		3	10	80	4	2	0.80	0.40
88250	2	Listening	MC	Core	1	924		5	82	3	8	2	0.82	0.44
88246	3	Listening	MC	Core	1	924		16	64	10	7	2	0.64	0.38
88248	4	Listening	MC	Core	1	924		5	87	2	3	2	0.87	0.47
88249	5	Listening	MC	Core	1	924		77	2	18	1	2	0.77	0.43
8061001	6	Listening	MC	Core	1	924		23	69	3	3	2	0.69	0.34
8061002	7	Listening	MC	Core	1	924		9	7	79	2	2	0.79	0.40
8061003	8	Listening	MC	Core	1	924		11	16	2	68	2	0.68	0.35
8061004	9	Listening	MC	Core	1	924		6	8	68	16	2	0.68	0.31
8062001	10	Listening	MC	Core	1	924		7	12	13	65	2	0.65	0.44
8062002	11	Listening	MC	Core	1	924		10	72	5	10	2	0.72	0.39
8062003	12	Listening	MC	Core	1	924		76	10	6	6	2	0.76	0.48
8062004	13	Listening	MC	Core	1	924		9	24	6	58	2	0.58	0.25
8029002	14	Listening	MC	Core	1	924		61	18	10	9	2	0.61	0.34
8029004	15	Listening	MC	Core	1	924		7	9	74	8	2	0.74	0.44
8029001	16	Listening	MC	Core	1	924		14	26	17	40	2	0.40	0.29
8029005	17	Listening	MC	Core	1	924		18	22	40	18	2	0.40	0.19
8029003	18	Listening	MC	Core	1	924		15	51	17	15	2	0.51	0.31
88239	1	Speaking	CR	Core	1	924	3	92				4	0.92	0.36
88253	2	Speaking	CR	Core	1	924	2	94				4	0.94	0.28
88440	3	Speaking	CR	Core	1	924	10	85				4	0.85	0.40
88254	4	Speaking	CR	Core	1	924	4	92				4	0.92	0.29
88385	5	Speaking	CR	Core	1	924	9	88				4	0.88	0.24
88255	6	Speaking	CR	Core	1	924	17	79				4	0.79	0.28
88258	7	Speaking	CR	Core	2	924	1	11	85			4	0.90	0.41
88441	8	Speaking	CR	Core	2	924	3	37	54			7	0.72	0.42
88442	9	Speaking	CR	Core	4	924	1	6	20	35	34	5	0.71	0.45
88244	10	Speaking	CR	Core	4	924	2	8	19	32	33	6	0.68	0.46
8065006	1	Reading	MC	Core	1	924		3	4	90	2	1	0.90	0.37
8065002	2	Reading	MC	Core	1	924		2	94	2	1	1	0.94	0.41
8065003	3	Reading	MC	Core	1	924		37	34	17	11	1	0.34	0.13
8065004	4	Reading	MC	Core	1	924		17	10	7	64	1	0.64	0.42
8065005	5	Reading	MC	Core	1	924		86	4	3	6	1	0.86	0.36
88279	6	Reading	MC	Core	1	924		79	16	2	2	1	0.79	0.36
8036001	7	Reading	MC	Core	1	924		1	24	72	2	1	0.72	0.31
8036003	8	Reading	MC	Core	1	924		1	94	3	1	1	0.94	0.35
8036002	9	Reading	MC	Core	1	924		56	16	10	17	1	0.56	0.22
8036004	10	Reading	MC	Core	1	924		29	28	12	31	1	0.31	0.22
8034005	11	Reading	MC	Core	1	924		4	9	62	24	1	0.62	0.16
8034002	12	Reading	MC	Core	1	924		10	72	7	10	1	0.72	0.42
8034003	13	Reading	MC	Core	1	924		2	4	9	83	1	0.83	0.45
8034004	14	Reading	MC	Core	1	924		18	53	17	10	1	0.53	0.36
8034001	15	Reading	CR	Core	2	924	41	28	27			4	0.41	0.51
8035003	16	Reading	MC	Core	1	924		7	5	13	72	2	0.72	0.42
8035001	17	Reading	MC	Core	1	924		61	20	12	5	2	0.61	0.36

Appendix A: Item Difficulty and Discrimination data (Continued)

Table A.9. Grades 9-12: Form E2 (Continued)

Item ID	Seq #	Modality	Item Type	Item Status	Max. Point	N-count	0	1/A	2/B	3/C	4/D	Omit	P-Value	Point Biserial
8035004	18	Reading	MC	Core	1	924		57	17	16	7	2	0.57	0.38
8035005	19	Reading	MC	Core	1	924		20	15	50	13	2	0.50	0.29
88393	1	Writing	MC	Core	1	924		2	8	87	2	1	0.87	0.39
88268	2	Writing	MC	Core	1	924		76	8	12	2	1	0.76	0.28
88264	3	Writing	MC	Core	1	924		12	78	3	5	2	0.78	0.23
88270	4	Writing	MC	Core	1	924		85	8	2	3	1	0.85	0.37
88269	5	Writing	MC	Core	1	924		17	19	61	1	2	0.61	0.29
8037001	6	Writing	MC	Core	1	924		71	15	7	5	2	0.71	0.45
8037006	7	Writing	MC	Core	1	924		1	6	76	15	2	0.76	0.39
8037003	8	Writing	MC	Core	1	924		4	83	7	4	2	0.83	0.45
8037004	9	Writing	MC	Core	1	924		7	58	6	27	2	0.58	0.31
8037005	10	Writing	MC	Core	1	924		13	4	14	68	1	0.68	0.31
88278	11	Writing	CR	Core	2	924	35	55	8			3	0.35	0.30
88276	12	Writing	CR	Core	4	924	8	22	45	17	4	3	0.45	0.56
88447	13	Writing	CR	Core	4	924	19	33	31	10	2	5	0.33	0.43

MontCAS

(Montana Comprehensive Assessment System)

English Language Proficiency Assessment

2008-2009

Score Reports Interpretation Guide

Denise Juneau, Superintendent
Montana Office of Public Instruction
www.opi.mt.gov

Copyright ©2009, held by the Montana Office of Public Instruction. All rights reserved. Printed in the U.S.A.
This assessment and its contents are the exclusive property of the State of Montana.

Contents

5 **Overview**

8 **Understanding the Individual Student Report**

10 **Understanding the Parent Report**

11 **Understanding the School Roster Report**

12 **Understanding the Summary Report**

13 **Understanding the Growth Report**

14 **Using MontCAS ELP Results**

Overview

The purpose of this guide is to assist educators and other stakeholders with understanding, interpreting, and using the results of the Montana English Language Proficiency Assessment. The MontCAS ELP is administered statewide to all Limited English Proficient (LEP) students.

The guide includes information on

- how and why the MontCAS ELP was developed,
- how the assessments are designed,
- how student performance is scored,
- how performance standards were determined,
- how assessment results are reported, and
- how results can be used to improve programs, instruction, and student performance.

Purpose of the MontCAS ELP. The annual assessment of LEP students in Montana fulfills a requirement of the No Child Left Behind Act of 2001. One objective is to measure individual student's progress in achieving proficiency in speaking, listening to, comprehending, reading, and writing English. A second objective is to measure the success of language development programs in achieving adequate student growth in English proficiency in districts participating in Title III.

Development of the MontCAS ELP. The MontCAS ELP is an edited version of the English Language Proficiency test developed for the Mountain West Consortium, of which Montana was a member. The first administration of the MontCAS ELP occurred in the fall of 2006. In February 2007, a panel of Montana educators met to set standards for the MontCAS ELP in the form of cut scores for each proficiency level by grade. A second set of MontCAS ELP forms was developed and administered in 2007. Those forms, which were similar in structure to the previous forms but included some new items, were equated to the 2006 forms. Thus scores in 2007 were reported on the same scale as 2006 and the same proficiency level cut scores applied. MontCAS ELP forms were changed somewhat for the 2008 administration.

This year's forms, which were built from items that appeared on the 2006 and 2007 forms, were shorter in terms of number of points per language domain than their predecessors. This shortening was achieved by selecting items for level 1 and level 2 forms that were more appropriate to the abilities of students administered at each level. As in 2007, though, the 2008 forms were equated to previously administered forms so that results are reported on the same scale and the proficiency level cut scores are the same as in 2006 and 2007.

Structure of the MontCAS ELP. The MontCAS ELP is comprised of tests in four domains–Listening, Speaking, Reading, and Writing. Scores are reported for each of these domains, as well as for Comprehension. The Comprehension score is calculated using a subset of Listening and Reading items.

The MontCAS ELP is administered by grade span.

Grade Span	Form
K	A
1-2	B1 or B2
3-5	C1 or C2
6-8	D1 or D2
9-12	E1 or E2

In all grade spans, except for K, there are two separate test forms, a Level 1 form intended for Beginning students and a Level 2 form intended for more proficient students. Having separate forms centered on two different ability levels made it possible to

shorten the forms. Note that no "mixed" scores can be reported: if, for example, a student took both B1 and B2 test forms, results have been reported for only one form.

Reported Scores. Student performance in each of the five language domains and on the overall (Total MontCAS ELP) test is reported in terms of raw score, scaled score, and proficiency level.

Raw Scores. The raw score is the total number of correct answers on multiple-choice items plus the number of points earned on open-ended items. Raw scores on the MontCAS ELP can only be compared for the same domain and the same test form. For example, a Form B1 raw score cannot be compared to a Form B2 raw score.

Note: The Writing raw score for (Kindergarten level) Form A was calculated as follows: 1 point was allocated for each skill on the Writing Checklist that the student "does most of the time" or of which they "demonstrate mastery." Thus, the Writing Checklist generated a maximum raw score of 22 points.

Scaled Scores. Scaled scores are derived from raw scores and provide results for alternate forms (e.g., B1 and B2) on a common scale. MontCAS ELP scaled scores can be compared for the same domain and the same grade-span test (A, B, C, D or E). For example, all Form C Reading scaled scores can be compared, regardless of whether the student took the C1 or the C2 Reading test. However, Form C scaled scores cannot be compared to Form D scaled scores.

Total MontCAS ELP Proficiency Levels. For the total score, four proficiency levels are reported: Novice (N), Nearing Proficiency (NP), Proficient (P), and Advanced (A). These are based on the total scaled score and provide a holistic estimate of the student's English proficiency. It is important to note that students at the same overall Proficiency Level may have different profiles of competence across the language domains.

Domain Proficiency Levels. Within each domain, two proficiency levels are reported, based on the student's scaled score: Below Proficient (BP) and Proficient or Above (PA). (Individual language domain tests are not long enough to reliably provide more than two levels of proficiency.)

Incomplete Testing. Students were required to take all four language domain tests. If a student did not take one or more of the domain tests, the reports will show dashes in place of scores for that domain. The reported Total MontCAS ELP score is based on the domain tests for which there are scores. Thus, if a student failed to take the Speaking Test for whatever reason, the Total MontCAS ELP score will be based on a raw score of zero in Speaking. The reported Comprehension scores—which are based on a subset of Listening and Reading scores—will be affected in the same way if the student failed to take either the Listening or Reading Test.

Cut Scores. The table below shows the MontCAS ELP Total scaled score range that corresponds to each proficiency level. Within a grade cluster (e.g., 3-5), cut scores may vary across each grade. Scaled scores should not be compared across grade clusters (e.g., 1-2 versus 3-5) but can be compared within a grade cluster. In those grade clusters with level 1 and 2 forms, the cut scores in each grade are the same regardless of the form administered.

		Scaled Score Range for Proficiency Levels			
Forms	Grade	Novice (N)	Nearing Proficiency (NP)	Proficient (P)	Advanced (A)
A	K	Below 363	363-395	396-424	At or Above 425
B1/B2	1	Below 345	345-373	374-420	At or Above 421
	2	Below 373	373-407	408-465	At or Above 466
C1/C2	3	Below 361	361-383	384-416	At or Above 417
	4	Below 374	374-396	397-429	At or Above 430
	5	Below 387	387-406	407-453	At or Above 454
D1/D2	6	Below 367	367-388	389-412	At or Above 413
	7	Below 367	367-391	392-419	At or Above 420
	8	Below 370	370-391	392-436	At or Above 437
E1/E2	9	Below 370	370-392	393-420	At or Above 421
	10	Below 373	373-395	396-423	At or Above 424
	11	Below 376	376-399	400-434	At or Above 435
	12	Below 376	376-399	400-434	At or Above 435

MontCAS
(Montana Comprehensive Assessment System)
English Language Proficiency Assessment

INDIVIDUAL STUDENT REPORT
English Language Proficiency (ELP) Assessment
2008 - 2009

Denise Juneau, Superintendent
Montana Office of Public Instruction
www.opi.mt.gov

Student	GAMMON, JUDE
School	ABC School
System	ABC System (9999)
Grade	4
1 Test Form	C2
2 State Student ID	123333789
Birth Date	05/14/1999
Gender	M
Test Date	Fall 2008

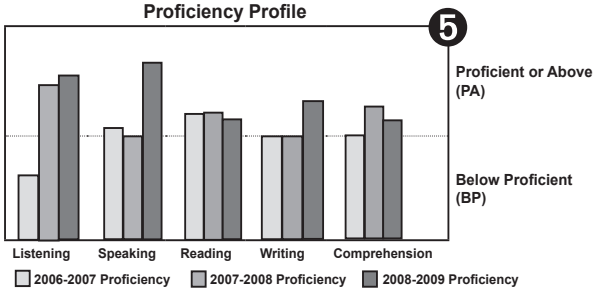
The NCLB Act of 2001 requires an annual assessment of English language proficiency for students identified as limited English proficient (LEP). The purpose of the assessment is to measure students' progress in achieving proficiency in academic English. The MontCAS English Language Proficiency (ELP) Assessment measures proficiency in listening, speaking, reading, writing, and comprehension (domains). The comprehension score is a composite score based on the listening and reading sections.

Novice students are beginning to participate in oral and written interactions of learned information to socialize, produce, and obtain information.
Nearing Proficient students demonstrate partial mastery of oral and written interactions of learned information to socialize, produce, and obtain information.
Proficient students demonstrate competent skills in oral and written interactions of learned information to socialize, produce, and obtain information in order to participate in academic work.
Advanced students demonstrate exceptional skills in oral and written interactions of learned information to socialize, produce, and obtain information in order to participate in academic work.

2007 - 2008	3 2008 - 2009 Total MontCAS ELP		
Proficiency Level	*Raw Score* (Max RS=72)	*Scaled Score*	*Proficiency Level* 6
Proficient (P)	57	426	Proficient (P)
	State Average Scaled Score	379.2	

2007 - 2008	2008 - 2009 Score Summary				
Proficiency Level		*Test*	*Raw Score*	*Scaled Score*	*Proficiency Level*
PA	**L**	Listening (Max RS=18)	15	115	PA
PA	**S**	Speaking (Max RS=18)	15	104	PA
PA	**R**	Reading (Max RS=18)	12	107	PA
PA	**W**	Writing (Max RS=18)	15	127	PA
PA	**C**	Comprehension (Max RS=36)	27	110	PA

Proficiency Profile 5



Legend: **RS**: Raw Score; **Max RS**: Maximum Possible Raw Score; **SS**: Scaled Score; **--** indicates test not taken **BP** = Below Proficient **PA** = Proficient or Above

Run Date: mm/ddy/yy

1 Test Form. Test forms are identified by a letter-number combination. The letter (A, B, C, D, or E) specifies the grade-span; the number specifies the difficulty level of the form (1 is for LEP students with beginner or novice skills in English; 2 is for the more proficient students). The exception is grade K (Form A), which does not have separate ability-level forms.

2 State Student ID. The state student ID is a unique number that is assigned to every student who receives educational services from a public school in Montana. This number follows the student from school to school throughout his or her K-12 career. The ID consists of 9 randomly generated digits, with no leading zeros.

3 **The Raw Score** is the total number of correct answers on multiple-choice items plus the number of points earned on open-ended items. A raw score can only be interpreted within the context of a given test form. Raw scores cannot be used to compare performance on different test forms. Scaled scores or scores derived from scaled scores should be used for those comparisons.

4 **Scaled Scores** are derived from raw scores and provide results for alternate forms (e.g., Forms B1 and B2) on a common scale. Scaled scores can be used to make comparisons among students and over time. However, scaled scores cannot be compared across test levels (e.g., B vs. C), or across different tests (e.g., Listening vs. Reading). To compare across different test levels, scaled scores must be converted to Proficiency Levels.

5 **The Proficiency Profile** summarizes ability across the language domains as well as growth from one year to the next, if a student has taken the MontCAS ELP for at least two years. The height of the bars shows how ability differs by language domain. The dotted line in the middle of the Proficiency Profile chart marks the cut score between the Below Proficient (BP) and the Proficient or Above (PA) levels, allowing you to see where student ability falls with respect to this criterion.

6 **Proficiency Levels** provide a holistic estimate of the student's English proficiency.

In general terms, the levels are:

Novice (N) – Students are beginning to participate in oral and written interactions of learned information to socialize, produce, and obtain information.

Nearing Proficiency (NP) – Students demonstrate partial mastery of oral and written interactions of learned information to socialize, produce, and obtain information.

Proficient (P) – Students demonstrate competent skills in oral and written interactions of learned information to socialize, produce, and obtain information in order to participate in academic work.

Advanced (A) – Students demonstrate exceptional skills in oral and written interactions of learned information to socialize, produce, and obtain information in order to participate in academic work.

The results of your student's English Language Proficiency Assessment are shown in this report by raw score, scaled score and performance level.

Raw score refers to the number of points a student has earned for a particular test. Raw scores should not be compared across language domains. A maximum raw score is shown for each language domain and the Total MontCAS.

Scaled scores are derived from raw scores and permit comparisons between level 1 and 2 forms (e.g., Form C1 and C2) within a grade cluster. Scaled scores range from 0 to 200.

Performance levels describe a student's performance on the MontCAS ELP assessment and are based on the total scaled score. The MontCAS ELP reports four performance levels for the total score (N, NP, P, A), which are organized into two groups for each domain (BP, PA). These performance levels are described in more detail on the back cover.

YOUR STUDENT'S RESULTS

The following charts show your student's performance on the English Language Proficiency Assessment. These charts include raw scores, scaled scores, and performance levels.

Total MontCAS ELP. This table indicates your student's overall performance on the 2008 - 2009 assessment. For comparative purposes, your student's overall proficiency level for last year, 2007 - 2008, and average state results for the current year are included. The score summary and proficiency profile on the next page provide more detailed information about how your child performed in each language domain.

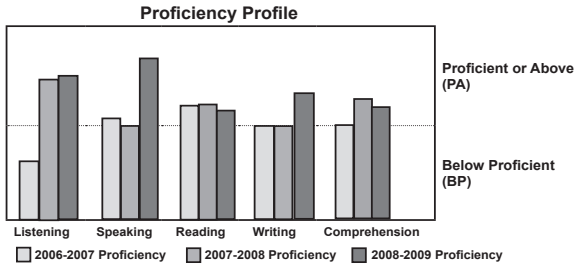
2007 - 2008	2008 - 2009 Total MontCAS ELP		
Proficiency Level	*Raw Score* (Max RS=72)	*Scaled Score*	*Proficiency Level*
Proficient (P)	57	426	Proficient (P)
	State Average Scaled Score	379.2	

Page 2

Score Summary. The Score Summary chart provides your student's results for each of five components of the ELP assessment: Listening, Speaking, Reading, Writing and Comprehension. The maximum raw score (Max RS) is indicated for each component. For example, the maximum raw score (Max RS) that could be earned for the Listening test was 18 points.

2007 - 2008	2008 - 2009 Score Summary				
Proficiency Level	*Test*		*Raw Score*	*Scaled Score*	*Proficiency Level*
PA	L	Listening (Max RS=18)	15	115	PA
PA	S	Speaking (Max RS=18)	15	104	PA
PA	R	Reading (Max RS=18)	12	107	PA
PA	W	Writing (Max RS=18)	15	127	PA
PA	C	Comprehension (Max RS=36)	27	110	PA

Proficiency Profile. The profile indicates your student's performance across the language domains, as well as growth from one year to the next.



Legend: **RS**: Raw Score; **Max RS**: Maximum Possible Raw Score; **SS**: Scaled Score; **--** indicates test not taken
BP = Below Proficient **PA** = Proficient or Above

Page 3

A customized parent report was generated for each LEP student who participated in the fall 2008 MontCAS English Language Proficiency (ELP) Assessment. This report was based on the school-level individual student report and should be shared by classroom teachers during parent-teacher conferences or other interactions with parents. The report includes detailed results of a student's ELP test performance, including raw scores, scaled scores and performance levels, in each language domain and for the total MontCAS ELP. The proficiency profile permits a comparison of student ability across the language domains and in comparison to average performance across the state.

Section A provides an explanation of terms—raw score, scaled scores, and performance levels—used in the Parent Report.

Section B shows the student's overall performance on the assessment in the Total MontCAS ELP table. The student's total raw score, scaled score, and proficiency level are provided, along with the Average State Scaled Score for this grade, for comparison.

Section C provides more detailed information about student performance in the Score Summary chart. The chart shows student results for each component of the ELP assessment: Listening, Speaking, Reading, Writing and Comprehension. The raw score, scaled score, and proficiency levels are listed for each of the five components.

Section D illustrates student performance in relation to the proficiency levels for up to 3 years (2006-2007, 2007-2008, 2008-2009), if a student took the MontCAS ELP assessment more than one year. The Proficiency Profile chart shows the scaled score "cut" line between proficiency levels Below Proficient (BP) and Proficient or Above (PA).

MontCAS
(Montana Comprehensive Assessment System)
English Language
Proficiency Assessment

CONFIDENTIAL
SCHOOL ROSTER
English Language Proficiency (ELP) Assessment
Grade 4
A 2008 - 2009
ABC School

Denise Juneau, Superintendent
Montana Office of Public Instruction
www.opi.mt.gov

SYSTEM: **ABC System (9999)** Test Date: Fall 2008

B Student Name / Number of Students Listed: 10		Gender	Test Form	C Listening			Speaking			Reading			Writing			Comprehension			D Total		
				RS	SS	Prof	RS	SS	Prof	RS	SS	Prof	RS	SS	Prof	RS	SS	Prof	RS	SS	Proficiency Level
BUTT, LEONEL D.	DOB: 01/21/1999	M	C2	13	107	PA	11	92	BP	10	101	PA	4	86	BP	21	101	PA	38	393	Nearing Proficiency
State ID#: 123456789																					
DESAI, BAILEE	DOB: 04/05/1999	F	C2	8	92	BP	9	87	BP	5	86	BP	2	77	BP	10	84	BP	24	372	Novice
State ID#: 123467890																					
EDENS, WARREN	DOB: 07/21/1998	M	C2	15	115	PA	16	105	PA	14	113	PA	7	96	BP	25	107	PA	52	416	Proficient
State ID#: 235678907																					
FANCHER, ELAINA L.	DOB: 08/01/1998	F	C2	16	121	PA	18	133	PA	15	117	PA	13	117	PA	27	110	PA	62	440	Advanced
State ID#: 435621897																					
GAMMON, JUDE	DOB: 05/14/1999	M	C2	15	115	PA	15	104	PA	12	107	PA	15	127	PA	27	110	PA	57	426	Proficient
State ID#: 860847350																					
HARRIMAN, MOHAMMAD M.	DOB: 10/20/1998	M	C2	17	129	PA	16	117	PA	15	117	PA	15	127	PA	30	116	PA	63	443	Advanced
State ID#: 975089899																					
HOCHSTETLER, SONIA D.	DOB: 09/22/1998	F	C2	9	95	BP	15	104	PA	6	90	BP	2	77	BP	10	84	BP	32	384	Nearing Proficiency
State ID#: 775534221																					
KEHOE, JAYDON J.	DOB: 06/12/1999	M	C2	13	107	PA	15	104	PA	12	107	PA	8	99	BP	21	101	PA	48	409	Proficient
State ID#: 896453311																					
KINGSTON, RYAN J.	DOB: 06/11/1999	M	C2	17	129	PA	16	109	PA	16	122	PA	12	113	PA	28	112	PA	61	437	Proficient
State ID#: 353243678																					
KIPP, ISAIAS D.	DOB; 05/31/1999	M	C2	14	111	PA	13	97	BP	10	101	PA	9	102	PA	22	102	PA	46	405	Proficient
State ID#: 796685767																					

Legend: **RS:** Raw Score; **Max RS:** Maximum Possible Raw Score; **SS**: Scale Score; **--** indicates test not taken **BP** = Below Proficient **PA** = Proficient or Above Page 1
Note: Any students who took the assessment with non-standard accommodations are marked with † symbol.
Run Date: mm/dd/yy

The MontCAS ELP School Roster report lists all students—in a single school in a single grade—who took the MontCAS ELP in a certain year. The School Roster report includes the following information:

Section A shows the grade, the assessment year, the school name, and system name.

Section B lists each student alphabetically, along with his or her state student ID number, date of birth, and gender. The Test Form column identifies the specific test form administered to the students.

Section C lists each student's raw score (RS), scaled score (SS), and proficiency level (Prof), in each language domain (Speaking, Listening, Reading, Writing, and Comprehension). Note that the Comprehension score is based on a subset of items from the Listening and Reading sections of the assessment. The language domain proficiency levels are: Below Proficient (BP) and Proficient or Above (PA).

Section D lists each student's Total MontCAS ELP raw score, total scaled score, and proficiency level: Novice (N), Nearing Proficiency (NP), Proficient (P), or Advanced (A).

MontCAS
(Montana Comprehensive Assessment System)
English Language
Proficiency Assessment

SYSTEM SUMMARY REPORT
English Language Proficiency (ELP) Assessment
A **Grade 4**
2008 - 2009

Denise Juneau, Superintendent
Montana Office of Public Instruction
www.opi.mt.gov

SYSTEM: **ABC System (9999)**

Test Form: C**1, C2**
Test Date: **Fall 2008**

B

Proficiency Level	Listening		Speaking		Reading		Writing		Comprehension	
	Scaled Score Range	Number and Percent of Students	Scaled Score Range	Number and Percent of Students	Scaled Score Range	Number and Percent of Students	Scaled Score Range	Number and Percent of Students	Scaled Score Range	Number and Percent of Students
Proficient or Above (PA)	At or Above 99	7 (70%)	At or Above 100	7 (70%)	At or Above 100	8 (80%)	At or Above 100	5 (50%)	At or Above 100	8 (80%)
Below Proficient (BP)	Below 99	3 (30%)	Below 100	3 (30%)	Below 100	2 (20%)	Below 100	5 (50%)	Below 100	2 (20%)
D N Students:	10*		10		10*		10*		10*	
Mean Scaled Score: System:	113.4		118.4		114.3		114.4		114.6	
Mean Scaled Score: State:	108.0		112.9		101.8		91.1		103.9	
Median Scaled Score: System:	116		118		114		123		114	
Median Scaled Score: State:	107		110		105		94		104	

*Summary statistics exclude students who did not take this subtest.

C

Proficiency Level	Total		
	Scaled Score Range	Number of Students	Percent
Advanced (A)	At or Above 430	2	20%
Proficient (P)	397 - 429	5	50%
Nearing Proficiency (NP)	374 - 396	2	20%
Novice (N)	Below 374	1	10%
N Students:		10	
Mean Scaled Score: System:		427.9	
Mean Scaled Score: State:		402.3	
Median Scaled Score: System:		430	
Median Scaled Score: State:		403	

Legend: **Mean Scaled Score:** The arithmetic average of a set of scaled scores. It is found by adding all the scores in the distribution and dividing by the total number of scores.
Median Scaled Score: The middle score in a distribution or set of ranked scaled scores. Half the scores in the set are below the median, and half are above it (the 50th percentile).

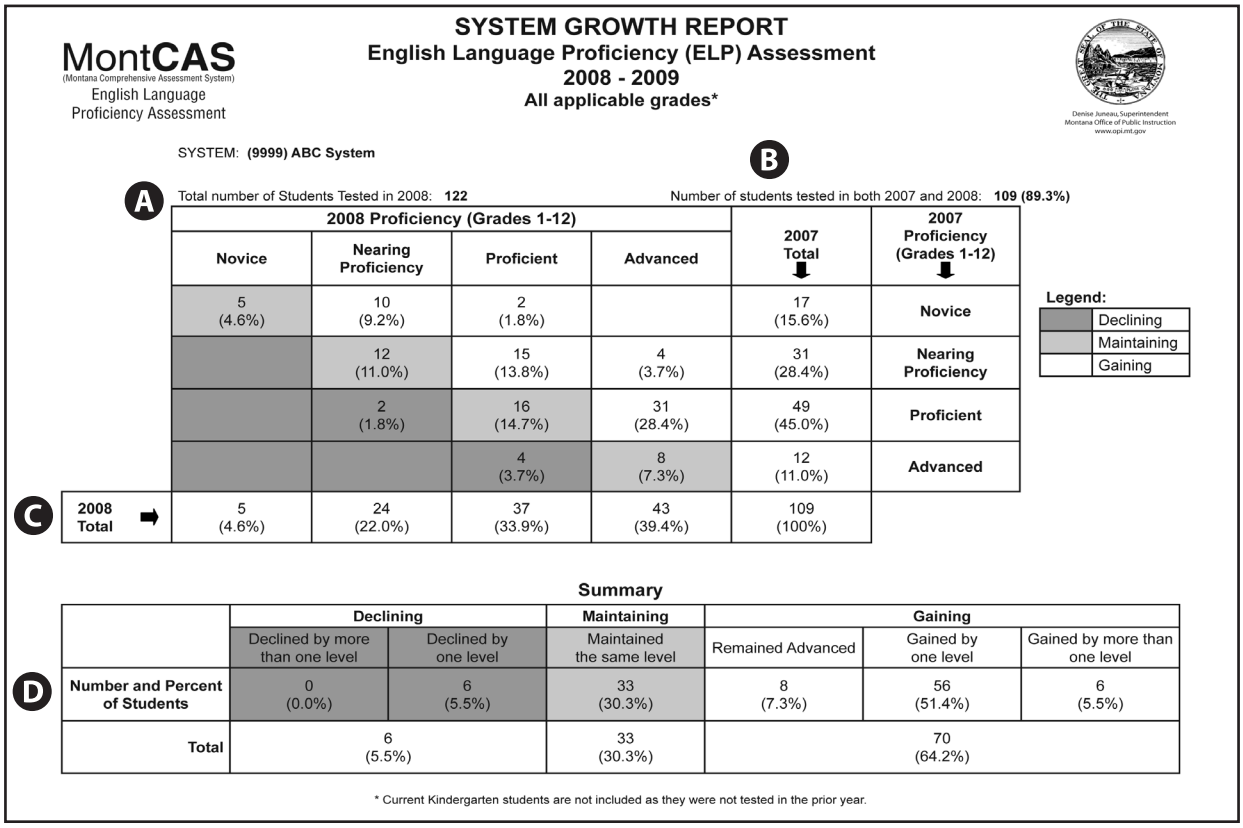
The MontCAS ELP System and School Summary Reports show the distribution of scores by grade within a system or school. The reports are produced even if the number of LEP students in a particular grade is very small. Reports for less than 10 students include a footer indicating that they may not be distributed to the public; the student information is protected by The Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99).

Section A shows the grade, the assessment year, and the system name.

Section B For each language domain (Speaking, Listening, Reading, Writing, and Comprehension), the report shows—in the Number and Percent of Students columns—the number and percent of students whose scores placed them in each of the two Proficiency Level groupings: Below Proficient (BP) and Proficient or Above (PA).

Section C The Total MontCAS ELP section shows scaled scores corresponding to each of 4 overall proficiency levels—Novice (N), Nearing Proficiency (NP), Proficient (P), and Advanced (A). The Number of Students column shows the number of students whose performance placed them in each category and the Percent column represents that number as a percentage of the students in this grade who were tested. For example, the 5 in the Proficient (P) cell of the sample report above indicates that 5 students in the system scored in the Proficient (P) range, which is 50% of the students in this grade.

Section D The N Students line shows the total number of students in the system in this grade for whom there is a language domain score and a total score. For example, the sample report shows that 10 4th-grade students took the Speaking Test. The Mean Scaled Score line shows the average scaled score in each domain and overall for all tested students in the system. For example, the sample report shows that the mean scaled score on the Speaking Test for this system was 118.4. The Median Scaled Score line shows the median scaled score in each domain and overall. The state mean and median are also shown for each domain and overall. Note that means and medians are shown only if N is 10 or greater.



SYSTEM GROWTH REPORT
English Language Proficiency (ELP) Assessment
2008 - 2009
All applicable grades*

MontCAS (Montana Comprehensive Assessment System) English Language Proficiency Assessment

Denise Juneau, Superintendent
Montana Office of Public Instruction
www.opi.mt.gov

SYSTEM: **(9999) ABC System**

A Total number of Students Tested in 2008: **122**

B Number of students tested in both 2007 and 2008: **109 (89.3%)**

	2008 Proficiency (Grades 1-12)				2007 Total	2007 Proficiency (Grades 1-12)
	Novice	Nearing Proficiency	Proficient	Advanced		
	5 (4.6%)	10 (9.2%)	2 (1.8%)		17 (15.6%)	Novice
		12 (11.0%)	15 (13.8%)	4 (3.7%)	31 (28.4%)	Nearing Proficiency
		2 (1.8%)	16 (14.7%)	31 (28.4%)	49 (45.0%)	Proficient
			4 (3.7%)	8 (7.3%)	12 (11.0%)	Advanced
C 2008 Total	5 (4.6%)	24 (22.0%)	37 (33.9%)	43 (39.4%)	109 (100%)	

Legend: Declining; Maintaining; Gaining

Summary

	Declining		Maintaining	Gaining		
	Declined by more than one level	Declined by one level	Maintained the same level	Remained Advanced	Gained by one level	Gained by more than one level
D Number and Percent of Students	0 (0.0%)	6 (5.5%)	33 (30.3%)	8 (7.3%)	56 (51.4%)	6 (5.5%)
Total	6 (5.5%)		33 (30.3%)	70 (64.2%)		

* Current Kindergarten students are not included as they were not tested in the prior year.

The MontCAS ELP System Growth Report shows the proficiency level profile within a system for those students who were assessed with the MontCAS ELP in both 2007 and 2008 (and have been confirmed by a State ID # match). Please note that System Growth Reports are provided only when there are 10 or more students who were tested (and matched by State ID #) in both 2007-2008 and 2008-2009 MontCAS ELP assessments. If the system has fewer than 10 students, Individual Reports should be examined to determine growth. The Growth Report includes the following information:

Section A shows the system name and total number of students from the designated grade or grades tested in 2008. The sample report shows growth for grades 1-12. Kindergarten is not included in the sample because these students were not tested in the prior year.

Section B shows the total number (and percentage) of students assessed in 2008 and matched by State ID # to 2007.

Section C shows a distribution of students by proficiency level for both 2007 and 2008 and how the proficiency of students in 2007 changed in 2008. Student proficiency level in 2007 is shown in the rows and summarized in the second to the last column on the right. So, for example, 31 students (28.4%) performed at the Nearing Proficiency level and 49 students (45.0%) at the Proficiency level in 2007. Student proficiency level in 2008 is shown in the columns and summarized in the last row on the bottom. So, for example, 37 students (33.9%) performed at the Proficient level in 2008. Thus comparing the 2007 Total column to the bottom row (2008 Total) shows how the distribution of performance for these students changed from 2007 to 2008. Each cell in the table shows how the students at a particular level in 2007 changed in 2008. So, for example, of those 49 students (middle row) who performed at the Proficient level in 2007, 16 (14.7%) tested at Proficient in 2008, and 28.4% tested at Advanced. The cells on the diagonal (upper left to lower right) show students whose proficiency level did not change. Those below the diagonal declined one or more levels from 2007 to 2008 and those above the diagonal gained one or more levels from 2007 to 2008.

Section D summarizes the changes from 2007 to 2008 shown in the upper panel. The bottom row aggregates students according to how their level changed and categorizes them as declining, maintaining, or gaining. Students who tested at Advanced (A) in both 2007 and 2008 were counted in the 'gaining' category.

Using MontCAS ELP Results

Monitoring Progress. MontCAS ELP test results can be used to determine whether students are making progress in developing English proficiency overall and within each language domain. To make comparisons between one year and the next, proficiency levels should be used. (Note that within a grade span, scaled scores can also be compared from year to year, as long as the student is being assessed with the same-letter form. Scaled scores cannot be used to monitor progress from year to year when students have moved to the next grade span, that is, in 1st grade, 3rd grade, 6th grade, and 9th grade.)

Informing Instruction. MontCAS ELP test results can be used to design instruction that capitalizes on students' strengths and addresses their weaknesses. Proficiency levels provide useful information on an individual student's profile across the language domains. For example, two students may both score as Proficient overall but have different strengths and weaknesses in the language domains. One may be lagging behind in Speaking, the other in Reading. With this information, instruction can be tailored to the individual student's needs.

Montana's Definition of "Proficient" for LEP Students Who Participate in the English Language Proficiency Assessment

In order to determine when LEP students become proficient, districts will take into account multiple measures which include:

- A score of Proficient (P) or Advanced (A) overall on the ELP assessment along with a rating of Proficient or Above (PA) in all domains (Listening, Speaking, Reading, and Writing). Students scoring as Proficient (P) should demonstrate a proficient score on the ELP assessment for two consecutive years. Students scoring as Advanced (A), along with additional measures and teacher input, would be considered proficient and not expected to take the ELP assessment again.
- Input from additional measures of reading, writing, or language development available from school assessments that link to the district process in place for the identification of LEP students.

This recommendation is based on input from representative school district staff members that serve LEP students across the state, a review of practices in other states, and input from psychometricians.

Denise Juneau, Superintendent
Montana Office of Public Instruction
www.opi.mt.gov

QAI-04720